

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 16:39:56 ; Search time 76 Seconds  
(without alignments)  
59.645 Million cell updates/sec

Title: US-09-980-523A-6  
Perfect score: 124  
Sequence: 1 RREYDFAFRDLICIVYRDGNPY 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 2060471115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*
- 15: sp.virus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	124	100.0	99	12	Q919B2 human papill
2	124	100.0	130	12	Q919C8 human papill
3	124	100.0	130	12	Q919C2 human papill
4	124	100.0	130	12	Q919C0 human papill
5	124	100.0	130	12	Q919B8 human papill
6	124	100.0	130	12	Q919B4 human papill
7	124	100.0	138	12	Q919D2 human papill
8	124	100.0	143	12	Q919B6 human papill
9	124	100.0	151	12	O12335 human papill
10	124	100.0	151	12	O12336 human papill
11	124	100.0	151	12	O9WMP5 human papill
12	124	100.0	151	12	O9WMP4 human papill
13	124	100.0	151	12	O9WMP3 human papill
14	124	100.0	151	12	O9WMP2 human papill
15	124	100.0	151	12	Q89640 human papill
16	124	100.0	151	12	Q89887 human papill

17	124	100.0	151	12	Q89648 human papill
18	124	100.0	151	12	Q89708 human papill
19	124	100.0	151	12	Q89755 human papill
20	124	100.0	151	12	Q80966 human papill
21	124	100.0	151	12	Q89852 human papill
22	124	100.0	151	12	Q9W8C3 human papill
23	124	100.0	158	12	Q9QDH7 human papill
24	124	100.0	158	12	Q9QDH5 human papill
25	124	100.0	158	12	Q9QDH3 human papill
26	124	100.0	158	12	Q9QDH9 human papill
27	124	100.0	158	12	Q9WH13 human papill
28	124	100.0	158	12	Q8QRE1 human papill
29	124	100.0	158	12	Q8QRE0 human papill
30	124	100.0	158	12	Q8QRD8 human papill
31	124	100.0	158	12	Q8QRD7 human papill
32	124	100.0	158	12	Q8QRD6 human papill
33	124	100.0	158	12	Q8QRD5 human papill
34	124	100.0	158	12	Q8QHT0 human papill
35	124	100.0	158	12	Q8QHN0 human papill
36	124	100.0	161	12	Q919D4 human papill
37	124	100.0	161	12	Q919C6 human papill
38	124	100.0	161	12	Q919B1 human papill
39	124	100.0	161	12	Q919A9 human papill
40	123	99.2	143	12	Q919C4 human papill
41	121	97.6	158	12	Q8QRD9 human papill
42	120	96.8	151	12	Q9W931 human papill
43	119	96.0	158	12	Q8QHP5 human papill
44	117	94.4	130	12	Q919D0 human papill
45	116	93.5	151	12	Q80963 human papill

#### ALIGNMENTS

##### RESULT 1

Q919B2 ID Q919B2 PRELIMINARY; PRT; 99 AA.  
AC Q919B2  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC13;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF404704; AAL01365.1;  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
FT NON\_TER 1  
FT NON\_TER 99  
SQ SEQUENCE 99 AA; 12005 MW; C2B96025EC370E38 CRC64;

Query Match 100.0%; Score 124; DB 12; Length 99;  
Best Local Similarity 100.0%; Pred. No. 3.9e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

##### QY 1 RREYDFAFRDLICIVYRDGNPY 22

|||||  
15 RREYDFAFRDLICIVYRDGNPY 36

##### RESULT 2

Q919C8 ID Q919C8 PRELIMINARY; PRT; 130 AA.

AC Q919C8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC8;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
cervical cancer isolates from Australia and New Caledonia.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF404696; AAL01355.1; -.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 130 AA; 15778 MW; 2830147D378B0DC9 CRC64;  
  
Query Match 100.0%; Score 124; DB 12; Length 130;  
Best Local Similarity 100.0%; Pred. No. 5.1e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RREYDFAFRDLCIVYRDGNPY 22  
DB 18 RREYDFAFRDLCIVYRDGNPY 39  
|||||  
RESULT 3  
Q919C2 PRELIMINARY; PRT; 130 AA.  
AC Q919C2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC8;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
cervical cancer isolates from Australia and New Caledonia.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF404699; AAL01355.1; -.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 130 AA; 15779 MW; 26D0147D396B0929 CRC64;  
  
Query Match 100.0%; Score 124; DB 12; Length 130;  
Best Local Similarity 100.0%; Pred. No. 5.1e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RREYDFAFRDLCIVYRDGNPY 22  
DB 18 RREYDFAFRDLCIVYRDGNPY 39  
|||||  
RESULT 4  
Q919C0 PRELIMINARY; PRT; 130 AA.  
AC Q919C0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC9;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
cervical cancer isolates from Australia and New Caledonia.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF404700; AAL01357.1; -.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 130 AA; 15779 MW; 26D0147D396B0929 CRC64;  
  
Query Match 100.0%; Score 124; DB 12; Length 130;  
Best Local Similarity 100.0%; Pred. No. 5.1e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RREYDFAFRDLCIVYRDGNPY 22  
DB 18 RREYDFAFRDLCIVYRDGNPY 39  
|||||  
RESULT 5  
Q919B8 PRELIMINARY; PRT; 130 AA.  
AC Q919B8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC10;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
cervical cancer isolates from Australia and New Caledonia.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF404701; AAL01359.1; -.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 130 AA; 15775 MW; 92D3C07BF96B092F CRC64;  
  
Query Match 100.0%; Score 124; DB 12; Length 130;  
Best Local Similarity 100.0%; Pred. No. 5.1e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RREYDFAFRDLCIVYRDGNPY 22  
DB 18 RREYDFAFRDLCIVYRDGNPY 39  
|||||  
RESULT 6  
Q919B4 PRELIMINARY; PRT; 130 AA.  
AC Q919B4;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.

OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC12;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
RL cervical cancer isolates from Australia and New Caledonia.";  
DR EMBL: AF404703; AAL01363.1; -  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 130 AA; 15792 MW; B6C2147D227EEDC CRC64;

Query Match 100.0%; Score 124; DB 12; Length 130;  
Best Local Similarity 100.0%; Pred. No. 5.1e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
DB 18 RREYDFAFRDLCIVYRDGNPY 39

## RESULT 7

ID Q919D2 PRELIMINARY; PRT; 138 AA.  
AC Q919D2;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC3;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
RL cervical cancer isolates from Australia and New Caledonia.";  
DR EMBL: AF404694; AAL01345.1; -  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 138 AA; 16696 MW; 481E5AEA90895FC2 CRC64;

Query Match 100.0%; Score 124; DB 12; Length 138;  
Best Local Similarity 100.0%; Pred. No. 5.5e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
DB 26 RREYDFAFRDLCIVYRDGNPY 47

## RESULT 8

ID Q919B6 PRELIMINARY; PRT; 143 AA.  
AC Q919B6;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE E6 protein (Fragment).  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.

OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPV16E6CC11;  
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;  
RT "Sequence variation and physical state of human papillomavirus type 16  
RL cervical cancer isolates from Australia and New Caledonia.";  
DR EMBL: AF404702; AAL01361.1; -  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
FT NON\_TER 1  
SQ SEQUENCE 143 AA; 17272 MW; 071F14EE3E6BE2AC CRC64;

Query Match 100.0%; Score 124; DB 12; Length 143;  
Best Local Similarity 100.0%; Pred. No. 5.7e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
DB 31 RREYDFAFRDLCIVYRDGNPY 52

## RESULT 9

ID O12335 PRELIMINARY; PRT; 151 AA.  
AC O12335;  
DT 01-JUL-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE E6 protein.  
GN E6 protein.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PC-4;  
RX MEDLINE=97437474; PubMed=9292007;  
RA Tonesello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,  
RA Beth-Giraldo E., Giraldo G.;  
RT "Sequence variations and viral genomic state of human papillomavirus  
RL type 16 in penile carcinomas from Ugandan patients.";  
RL J. Gen. Virol. 78:2199-2208(1997).  
DR EMBL: AF003015; AAB70732.1; -  
DR InterPro: IPR001334; E6.  
DR Pfam: PF00518; E6; 1.  
SQ SEQUENCE 151 AA; 18238 MW; BEF32A8B016CC8B8 CRC64;

Query Match 100.0%; Score 124; DB 12; Length 151;  
Best Local Similarity 100.0%; Pred. No. 6e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
DB 39 RREYDFAFRDLCIVYRDGNPY 60

## RESULT 10

ID O12336 PRELIMINARY; PRT; 151 AA.  
AC O12336;  
DT 01-JUL-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE E6 protein.  
GN E6 protein.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PC-7;

```
RA MEDLINE=97437474; PubMed=9292007;
RA Torneoello M.L., Buonaguro F.M., Meglio A., Buonaguro L.,
RA Beth-Giraldo E., Giraldo G.;
RT "Sequence variations and viral genomic state of human papillomavirus
RT type 16 in penile carcinomas from Ugandan patients.";
RL J. Gen. Virol. 78:2199-2208(1997).
DR EMBL; AF003016; AAB/0733.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18206 MW; 51C12A8B149C6D8A CRC64;

Query Match 100.0%; Score 124; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGPNY 22
DQ 39 RREYDFAFRDLCIVYRDGPNY 60

RESULT 11
Q9WMP5 PRELIMINARY; PRT; 151 AA.
AC Q9WMP5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ET182G;
RX MEDLINE=20112892; PubMed=10644829;
RA van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT "Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotypes in cervical carcinogenesis.";
RL J. Gen. Virol. 81:317-325(2000).
DR EMBL; AJ242681; CAB45381.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18320 MW; 617D2D5FD93F8917 CRC64;

Query Match 100.0%; Score 124; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGPNY 22
DQ 39 RREYDFAFRDLCIVYRDGPNY 60

RESULT 12
Q9WMP4 PRELIMINARY; PRT; 151 AA.
AC Q9WMP4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EC410T;
RX MEDLINE=20112892; PubMed=10644829;
```

```
RA van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT "Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotypes in cervical carcinogenesis.";
RL J. Gen. Virol. 81:317-325(2000).
DR EMBL; AJ388060; CAB45112.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18387 MW; E2244784BEBAC02 CRC64;

Query Match 100.0%; Score 124; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGPNY 22
DQ 39 RREYDFAFRDLCIVYRDGPNY 60

RESULT 13
Q9WMP3 PRELIMINARY; PRT; 151 AA.
AC Q9WMP3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AA/AF;
RX MEDLINE=20112892; PubMed=10644829;
RA van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
RA Verheijen R.H., Helmerhorst T.J., Meijer C.J., Walboomers J.M.;
RT "Analysis of Human papillomavirus type 16 E6 variants in relation to
RT p53 codon 72 polymorphism genotypes in cervical carcinogenesis.";
RL J. Gen. Virol. 81:317-325(2000).
DR EMBL; AJ388065; CAB45122.1; -.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
SQ SEQUENCE 151 AA; 18412 MW; E794A494FD0E8209 CRC64;

Query Match 100.0%; Score 124; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGPNY 22
DQ 39 RREYDFAFRDLCIVYRDGPNY 60

RESULT 14
Q9WMP2 PRELIMINARY; PRT; 151 AA.
AC Q9WMP2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AFC779T;
RX MEDLINE=20112892; PubMed=10644829;
RA van Duin M., Snijders P.J., Vossen M.T., Klaassen E., Voorhorst F.,
```

RT "Analysis of Human papillomavirus type 16 E6 variants in relation to  
RT p53 codon 72 polymorphism genotypes in cervical carcinogenesis.";  
RT J. Gen. Virol. 81:317-325(2000).  
DR EMBL; AJ388067; CAB45126.1; -.

Query Match	100.00;	Score 124;	DB 12;	Length 151;
Best Local Similarity	100.00;	Pred. No. 6e-12;		
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

D6 39 BREYDEAFRDL CIVYRGNPY 60

Q89640	PRELIMINARY;	PRT;	151 AA.
ID	Q89640;		
AC	Q89640;		
DT	01-NOV-1996 (TEMBLrel. 01, Created)		
DT	01-NOV-1996 (TEMBLrel. 01, Last sequence update)		
DT	01-JUN-2001 (TEMBLrel. 17, Last annotation update)		
DE	Early transforming protein E6.		
GN	E6.		
OS	Human papillomavirus.		
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;		
OC	Papillomavirus.		
OX	NCBI_TaxID=10586;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=OR 7574;		
RX	MEDLINE=96079021; Pubmed=7494284;		
RA	Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,		
RA	Janison S.A.;		
RT	"Human papillomavirus type 16 variant lineages in United States		
RT	populations characterized by nucleotide sequence analysis of the E6,		
RT	L2, and L1 coding segments.";		
RL	J. Virol. 69:7743-7753(1995).		
RN	[2]		

RC	STRAIN-OR 7574;
RA	Farmer A.D.;
RA	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RL	
DR	EMBL; U34129; AAA91676.1; -;
DR	EMBL; U34115; AAA91662.1; -;
DR	EMBL; U34120; AAA91667.1; -;
DR	EMBL; U34124; AAA91671.1; -;
DR	InterPro: IPR001334; E6.
DR	Pfam: PF00518; E6; 1.
DR	SEQUENCE 151 AA; 18348 MW; FE3F2A2FCF0A6CB2 CRC64;
SQ	

Best Local Similarity 100.0%; Pred. No. 6e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRLCLIVYRDGNPY 22  
|||||  
Db 39 RREYDFAFRLCLIVYRDGNPY 60

Job time : 77 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 16:39:56 ; Search time 22 Seconds  
(without alignments)  
41.476 Million cell updates/sec

Title: US-09-980-523A-6

Perfect score: 124

Sequence: 1 RREYDFAFRLCIVIRGPNY 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	158	1 VE6_HPV16	P03126 human papill
2	94	75.8	158	1 VE6_HPV68	P54667 human papill
3	93	75.0	149	1 VE6_HPV33	P06427 human papill
4	93	75.0	149	1 VE6_HPV35	P06428 human papill
5	92	74.2	149	1 VE6_HPV58	P26555 human papill
6	90	72.6	148	1 VE6_HPV34	P36811 human papill
7	88	71.0	148	1 VE6_HPV52	P36814 human papill
8	83	66.9	158	1 VE6_HPV45	P21735 human papill
9	82	66.1	158	1 VE6_HPV70	P50804 human papill
10	79	63.7	154	1 VE6_HPV53	P36815 human papill
11	79	63.7	158	1 VE6_HPV39	P24835 human papill
12	78	62.9	158	1 VE6_HPVNE	P27962 human papill
13	73	58.9	151	1 VE6_HPV51	P26554 human papill
14	72	58.1	149	1 VE6_HPV31	P17386 human papill
15	71	57.3	158	1 VE6_HPV18	P06463 human papill
16	70	56.5	155	1 VE6_HPV43	P19709 human papill
17	68	54.8	153	1 VE6_HPV30	P22159 human papill
18	68	54.8	191	1 VE6_RHPV1	P22159 rhesus papil
19	66	53.2	150	1 VE6_HPV26	P36807 human papill
20	65	52.4	155	1 VE6_HPV56	P24836 human papill
21	58	46.8	150	1 VE6_HPV44	P19710 human papill
22	57	46.0	150	1 VE6_HPV6A	Q84291 human papill
23	56	45.2	144	1 VE6_HPV54	Q81018 human papill
24	56	45.2	150	1 VE6_HPV6B	P06462 human papill
25	56	45.2	154	1 VE6_HPV40	P36812 human papill
26	54	43.5	150	1 VE6_HPV11	P04019 human papill
27	53	42.7	142	1 VE6_HPV32	P36810 human papill
28	52	41.9	154	1 VE6_HPV07	P36800 human papill
29	51	41.1	146	1 VE6_HPV28	P50802 human papill
30	51	41.1	152	1 VE6_HPV03	P36799 human papill
31	50	40.3	322	1 RBSC_BACSU	P36948 bacillus su
32	49	39.5	148	1 VE6_HPV10	P36802 human papill
33	48	38.7	150	1 VE6_PCPV1	Q02270 pygmy chimp

34 48 38.7 153 1 VE6\_HPV57 P22158 human papill  
35 48 38.7 155 1 VE6\_HPV66 Q0955 human papill  
36 47 37.9 150 1 VE6\_HPV42 P27229 human papill  
37 47 37.9 153 1 VE6\_HPV2A P25484 human papill  
38 46 37.1 159 1 VE6\_HPV27 P36808 human papill  
39 46 37.1 653 1 RN17\_HUMAN Q9BXT8 homo sapien  
40 45.5 36.7 342 1 GBPR\_AZOB P52661 azospirillum  
41 45 36.3 148 1 VE6\_HPV29 P50803 human papill  
42 45 36.3 150 1 VE6\_HPV13 Q02269 human papill  
43 45 36.3 411 1 SNC3\_HUMAN Q92956 homo sapien  
44 45 36.3 482 1 IFT5\_HUMAN Q13325 homo sapien  
45 44 35.5 114 1 YHIT\_SYNY3 P73481 synechocyst

#### ALIGNMENTS

RESULT 1  
VE6\_HPV16  
ID VE6\_HPV16 STANDARD; PRT; 158 AA.  
AC P03126;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE E6 protein.  
GN E6.  
OS Human papillomavirus type 16.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10581;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85246220; PubMed=2990099;  
RA Seedorf K., Krammer G., Durst M., Suhai S., Roweckamp W.G.;  
RT "human papillomavirus type 16 DNA sequence."  
RL Virology 145:181-185(1985).  
RN [2]  
RP SEQUENCE OF 31-50 FROM N.A.  
RX MEDLINE=90218027; PubMed=2157796;  
RA Schneider-Maunoury S., Pehau-Arnaut G., Breitburd F., Orth G.;  
RT "Expression of the human papillomavirus type 16 genome in SK-v cells,  
a line derived from a vulvar intraepithelial neoplasia.";  
RL J. Gen. Virol. 71:809-817(1990).  
CC -!- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.  
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.  
CC -!- MISCELLANEOUS: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE  
OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; K02718; AAA46939.1; -;  
DR EMBL; D00735; BAA00632.1; -;  
DR PIR; A03682; W6WLHS.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;  
KW Oncogene.  
FT ZN\_FING 37 73 POTENTIAL.  
FT ZN\_FING 110 146 POTENTIAL.  
SQ SEQUENCE 158 AA; 19187 MW; 01FEF5ADCDB37EB CRC64;

Query Match 100.0%; Score 124; DB 1; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 RREYDFAFRDLCLIVYRDGNPY 22
    |||||
DB 46 RREYDFAFRDLCLIVYRDGNPY 67

RESULT 2
VE6_HPV68          STANDARD;          PRT; 158 AA.
AC P54667;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 68.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=45240;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060129; PubMed=8904450;
RA Longuet M., Beaudenon S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RT related to the potentially oncogenic HPV39."
RL J. Clin. Microbiol. 34:738-744(1996).
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12732; AAA46958.1; --
DR PIR; A03683; W6WL33.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
DR KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;
KW Oncogene.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
FT ZN_FING 149 AA; 17652 MW; EDDB7444F9C8B1AF CRC64;
SQ SEQUENCE 149 AA; 17652 MW; 17652 MW; EDDB7444F9C8B1AF CRC64;

Query Match 75.8%; Score 94; DB 1; Length 149;
Best Local Similarity 72.7%; Pred. No. 1.3e-07;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 RREYDFAFRDLCLIVYRDGNPY 22
    |||||
DB 39 RSEYDFAFRDLCLIVYRDGNPY 60

RESULT 4
VE6_HPV35          STANDARD;          PRT; 149 AA.
AC P27228;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10587;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types."
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92124753; PubMed=1310198;
RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
RT human papillomavirus type 35."
RL Virology 186:770-776(1992).
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL
CC OF THIS VIRUS (ASSOCIATED WITH CANCER OF THE UTERINE CERVIX).
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67160; CAA47632.1; --
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
DR KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT ZN_FING 32 68 POTENTIAL.
FT ZN_FING 105 141 POTENTIAL.
FT ZN_FING 158 AA; 18796 MW; 46B37939CFBA6596 CRC64;
SQ SEQUENCE 158 AA; 18796 MW; 46B37939CFBA6596 CRC64;

Query Match 75.8%; Score 94; DB 1; Length 158;
Best Local Similarity 72.7%; Pred. No. 1.3e-07;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 RREYDFAFRDLCLIVYRDGNPY 22
    |||||
DB 41 RTEYVEFAFRDLCLIVYRDGNPY 62

RESULT 3
VE6_HPV33          STANDARD;          PRT; 149 AA.
AC P06427;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 33.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10586;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86200464; PubMed=3009902;
RA Cole S.T., Strebeck R.E.;
RT "Genome organization and nucleotide sequence of human papillomavirus
RT type 33, which is associated with cervical cancer.";
```



```
-----
CC  EMBL: X74477; CAA52561.1; -.
DR  EMBL: M74117; AAA46966.1; -.
DR  PIR: E40824; W6WL35.
DR  PIR: S36521; S36521.
DR  InterPro: IPR001334; E6.
DR  Pfam: PF00518; E6; 1.
KW  Early protein; DNA-binding; Nuclear protein; Zinc-finger;
KW  Oncogene.
FT  ZN_FING 30 66 POTENTIAL.
FT  ZN_FING 103 139 POTENTIAL.
SQ  SEQUENCE 149 AA; 18045 MW; C605D19AF3935021 CRC64;

Query Match      75.0%; Score 93; DB 1; Length 149;
Best Local Similarity 77.3%; Pred. No. 1.8e-07;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 RREYDFAFRDLICIVYRDGNPY 22
    | | | | | | | | | | | | | | | |
Db  39 RSEYDFACYDLICIVYREGQPY 60

RESULT 5
VE6_HPV58
ID  VE6_HPV58      STANDARD;      PRT;      149 AA.
AC  P26555;
DT  01-AUG-1992 (Rel. 23, Created)
DT  01-AUG-1992 (Rel. 23, Last sequence update)
DE  E6 protein.
DE  E6.
OS  Human papillomavirus type 58.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10598;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92024102; PubMed=1656594;
RA  Kirii Y., Iwamoto S., Matsukura T.;
RT  "Human papillomavirus type 58 DNA sequence.";
RL  Virology 185:424-427(1991).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D90400; BAA31845.1; -.
DR  PIR: E36779; W6WL50.
DR  InterPro: IPR001334; E6.
DR  Pfam: PF00518; E6; 1.
KW  Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT  ZN_FING 30 66 POTENTIAL.
FT  ZN_FING 103 139 POTENTIAL.
SQ  SEQUENCE 149 AA; 17794 MW; 79B3DC95831B158 CRC64;

Query Match      74.2%; Score 92; DB 1; Length 149;
Best Local Similarity 77.3%; Pred. No. 2.5e-07;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 RREYDFAFRDLICIVYRDGNPY 22
    | | | | | | | | | | | | | | | |
Db  39 RSEYDFVAFDLRIVYRDGNPF 60

RESULT 6
VE6_HPV34
ID  VE6_HPV34      STANDARD;      PRT;      148 AA.
AC  P36811;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  01-OCT-1994 (Rel. 30, Last annotation update)
DE  E6 protein.
DE  E6.
OS  Human papillomavirus type 34.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10613;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94265501; PubMed=8205838;
RA  Delius H., Hofmann B.;
RT  "Primer-directed sequencing of human papillomavirus types.";
RL  Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
```

```
-----
ID  VE6_HPV34      STANDARD;      PRT;      148 AA.
AC  P36811;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  01-OCT-1994 (Rel. 30, Last annotation update)
DE  E6 protein.
DE  E6.
OS  Human papillomavirus type 34.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10613;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94265501; PubMed=8205838;
RA  Delius H., Hofmann B.;
RT  "Primer-directed sequencing of human papillomavirus types.";
RL  Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X74476; CAA52555.1; -.
DR  PIR: S36515; S36515.
DR  InterPro: IPR001334; E6.
DR  Pfam: PF00518; E6; 1.
KW  Early protein; DNA-binding; Nuclear protein; Zinc-finger.
FT  ZN_FING 31 67 POTENTIAL.
FT  ZN_FING 104 140 POTENTIAL.
SQ  SEQUENCE 148 AA; 17735 MW; E2FCC6E62E4AF0DA CRC64;

Query Match      72.6%; Score 90; DB 1; Length 148;
Best Local Similarity 76.2%; Pred. No. 5.1e-07;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 RREYDFAFRDLICIVYRDGNP 21
    | | | | | | | | | | | | | | | |
Db  40 RCEYDFIFRDLICVYRKGP 60

RESULT 7
VE6_HPV52
ID  VE6_HPV52      STANDARD;      PRT;      148 AA.
AC  P36814;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  E6 protein.
DE  E6.
OS  Human papillomavirus type 52.
OC  Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC  Papillomavirus.
OX  NCBI_TaxID=10618;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94265501; PubMed=8205838;
RA  Delius H., Hofmann B.;
RT  "Primer-directed sequencing of human papillomavirus types.";
RL  Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC  -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC  STRANDED DNA (IN VITRO).
CC  -1- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X74481; CRA52585.1; -  
DR PIR; S36573; S36573.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
FT ZN\_FING 30 66  
FT ZN\_FING 103 139 POTENTIAL.  
SQ SEQUENCE 148 AA; 17898 MW; C5E9DCE341AB0DE CRC64;

Query Match 71.0%; Score 88; DB 1; Length 148;  
Best Local Similarity 77.3%; Pred. No. 1e-06; 5; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLICIVYRDGPNY 22  
Db 39 RREYVKELFTDLRIYVRDNNFY 60

## RESULT 8

VE6\_HPV45  
ID VE6\_HPV45 STANDARD; PRT; 158 AA.  
AC P21735;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE E6 protein.  
GN E6.  
OS Human papillomavirus type 45.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10593;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; PubMed=8205838;  
RA Delliuss H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kaplan J.B., Burk R.D.;  
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.  
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
CC STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X74479; CAA52573.1; -  
DR EMBL; M38198; AAA46973.1; -  
DR PIR; S36561; S36561.  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;  
KW Oncogene.  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 105 141 POTENTIAL.  
FT CONFLICT 10 40 R -> P (IN REF. 2).  
FT CONFLICT 30 30 I -> N (IN REF. 2).  
FT CONFLICT 118 118 R -> A (IN REF. 2).  
SQ SEQUENCE 158 AA; 18897 MW; F1CF10DD33AAA4C3E CRC64;

Query Match 66.9%; Score 83; DB 1; Length 158;  
Best Local Similarity 83.3%; Pred. No. 6.6e-06;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLICIVYRD 18  
Db 41 RTEYQFAFKDLICIVYRD 58

## RESULT 9

VE6\_HPV70  
ID VE6\_HPV70 STANDARD; PRT; 158 AA.  
AC P50804;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE E6 protein.  
GN E6.  
OS Human papillomavirus type 70.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=39457;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96249586; PubMed=8815087;  
RA Forslund O., Hansson B.G.;  
RT "Human papillomavirus type 70 genome cloned from overlapping PCR  
RT products: complete nucleotide sequence and genomic organization.";  
RL J. Clin. Microbiol. 34:802-809(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97060129; PubMed=8904450;  
RA Longuet M., Beaudenon S., Orth G.;  
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,  
RT related to the potentially oncogenic HPV39.";  
RL J. Clin. Microbiol. 34:738-744(1996).  
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE  
CC STRANDED DNA (IN VITRO).  
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U21941; AAC54850.1; -  
DR EMBL; U22461; AAC54880.1; -  
DR InterPro; IPR001334; E6.  
DR Pfam; PF00518; E6; 1.  
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger.  
FT ZN\_FING 32 68 POTENTIAL.  
FT ZN\_FING 105 141 POTENTIAL.  
FT CONFLICT 100 100 N -> D (IN REF. 2).  
SQ SEQUENCE 158 AA; 18617 MW; 6B610800D923D6DE CRC64;

Query Match 66.1%; Score 82; DB 1; Length 158;  
Best Local Similarity 75.0%; Pred. No. 9.4e-06;  
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EYVDFAFRDLICIVYRDGPNY 22  
Db 43 EVIEFAFSDLFIVIRNGEPI 62

## RESULT 10

VE6\_HPV53  
ID VE6\_HPV53 STANDARD; PRT; 154 AA.  
AC P36815;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)



```
RESULT 13
VE6_HPV51          STANDARD;          PRT;   151 AA.
AC P26554;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 51.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10595;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91303675; PubMed=1649326;
RA Lungu O., Crum C.P., Silverstein S.J.;
RT "Biologic properties and nucleotide sequence analysis of human
papillomavirus type 51."
RL J. Virol. 65:4216-4225(1991).
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
STRANDED DNA (IN VITRO).
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; J04353; AAA46950.1;
DR PIR; A32444; W6WL31.
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;
Oncogene.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
SQ SEQUENCE 149 AA; 17713 MW; 61D2A86C362767D9 CRC64;

Query Match          58.1%; Score 72; DB 1; Length 149;
Best Local Similarity 70.0%; Pred. No. 0.00031;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 EYDFAFRDLCIVYRDGNPY 22
   ||||| ||||| ||
DB 41 EVLDFAFDTLTVYRDDTPH 60
   ||||| ||||| ||

RESULT 15
VE6_HPV18          STANDARD;          PRT;   158 AA.
AC P06463;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 18.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10582;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87283882; PubMed=3039146;
RA Cole S.T., Danos O.;
RT "Nucleotide sequence and comparative analysis of the human
papillomavirus type 18 genome. Phylogeny of papillomaviruses and
repeated structure of the E6 and E7 gene products."
RL J. Mol. Biol. 193:599-608(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306665; PubMed=3018129;
RA Matlashewski G., Banks L., Wu-Liao J., Spence P., Pim D., Crawford L.;
RT "The expression of human papillomavirus type 18 E6 protein in
bacteria and the production of anti-E6 antibodies."
RL J. Gen. Virol. 67:1909-1916(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88188247; PubMed=2833614;
RA Inagaki Y., Tsunokawa Y., Takebe N., Nawa H., Nakanishi S.,
RA Terada M., Sugimura T.;
RT "Nucleotide sequences of cDNAs for human papillomavirus type 18
transcripts in HeLa cells."
RL J. Virol. 62:1640-1646(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053870; PubMed=3023067;
RA Schneider-Gaedcke A., Schwarz E.;
RT "Different human cervical carcinoma cell lines show similar
transcription patterns of human papillomavirus type 18 early genes."
RL EMBO J. 5:2285-2292(1986).
```

```
RN SEQUENCE FROM N.A.
RP MEDLINE=87218459; PubMed=3034571;
RA Seedorf K., Oltersdorf T., Kraemer G., Roewkamp W.;
RT "Identification of early proteins of the human papilloma viruses type
RT 16 (HPV 16) and type 18 (HPV 18) in cervical carcinoma cells.";
RL EMBO J. 6:139-144(1987).
RN [6]
RP ZINC-BINDING.
RX MEDLINE=89385606; PubMed=2550872;
RA Grossman S.R., Laimins L.A.;
RT "E6 protein of human papillomavirus type 18 binds zinc.";
RL Oncogene 4:1089-1093(1989).
CC -!- FUNCTION: THIS PROTEIN HAS TRANSFORMING ACTIVITY IN VITRO.
CC -!- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -!- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04354; CAA27879.1; -
DR EMBL; X05015; CAA28664.1; -
DR EMBL; M20325; AAA99514.1; -
DR EMBL; M26798; AAA46946.1; -
DR EMBL; X04773; CAA28466.1; -
DR EMBL; A06324; CAA00539.1; -
DR EMBL; A06328; CAA00542.1; -
DR PIR; G26251; W6WL18
DR InterPro; IPR001334; E6.
DR Pfam; PF00518; E6; 1.
KW Early protein; DNA-binding; Nuclear protein; Zinc-finger;
KW Oncogene.
FT ZN_FING 32 68 POTENTIAL.
FT ZN_FING 105 141 POTENTIAL.
FT CONFLICT 22 22 N -> S (IN REF. 4).
SQ SEQUENCE 158 AA; 18871 MW; 5BCF13CF43D157FA CRC64;

Query Match 57.3%; Score 71; DB 1; Length 158;
Best Local Similarity 60.0%; Pred. No. 0 00047;
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 EVYDFAFRDLCIVYRDGNPY 22
   ||:|||||:|||||:
Db 43 EVFEFAFKDLFVYVRDSIPH 62
```

Search completed: July 3, 2003, 16:43:55  
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 6, 2003, 13:08:09 ; Search time 1415 Seconds

(without alignments)  
251.803 Million cell updates/sec

Title: US-09-980-523A-6  
Perfect score: 124  
Sequence: 1 RREYDFAFRLCIVIRGPNY 22

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlp  
-O=/cpn2.1/USPTO.spool/US09980523/runat\_03072003.164406.4239/app\_query.fasta\_1.199  
-DB=EST-QFMT-fastap SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09980523@cgn.1.1.2874\_erunat\_03072003.164406.4239 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hct.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hct.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	58	46.8	443	12	BF911801	BF911801 MR4-UT005
C 2	58	46.8	681	12	BG810071	BG810071 mgct002xk
C 3	56	45.2	520	10	AV862231	AV862231 AV862231
C 4	56	45.2	586	10	AV868962	AV868962 AV868962
C 5	56	45.2	622	10	AV852103	AV852103 AV852103
C 6	55	44.4	758	17	AG166938	AG166938 Pan trogl
C 7	55	44.4	825	17	B20885	B20885 FLIE8-T7.2
C 8	55	44.4	890	10	BE543005	BE543005 601068977
C 9	55	44.4	966	17	CNS07ABD	AL436367 T3 end of
C 10	54	43.5	847	17	AZ741179	AZ741179 RPI1-24-7
C 11	54	43.5	873	12	BF106599	BF106599 601823239
C 12	53	42.7	551	10	BE067619	BE067619 MR4-BT035
C 13	53	42.7	644	10	BE067679	BE067679 MR4-BT035
C 14	53	42.7	755	17	AG029226	AG029226 Pan trogl
C 15	53	42.7	1044	10	BE539653	BE539653 601059444
C 16	52.5	42.3	821	12	BG645530	BG645530 EST507149
C 17	52	41.9	776	12	BF344794	BF344794 602014063
C 18	52	41.9	806	9	AI760789	AI760789 w16e07.x
C 19	52	41.9	916	12	BG111082	BG111082 602284775
C 20	52	41.9	955	17	CNS07AY1	AL437200 T3 end of
C 21	51	41.1	374	17	AQ002097	AQ002097 Cit-HSP-2
C 22	51	41.1	392	14	W61151	W61151 zd3lc08.s1
C 23	51	41.1	393	12	BF602565	BF602565 267827 MA
C 24	51	41.1	473	14	R80018	R80018 Y191a05.r1
C 25	51	41.1	543	17	TA91808P	AL459580 T. brucei
C 26	51	41.1	591	17	AZ969409	AZ969409 2M0242J10
C 27	51	41.1	591	17	AQ652922	AQ652922 Sheared D
C 28	51	41.1	596	13	BM630307	BM630307 170006875
C 29	51	41.1	676	17	BH280883	BH280883 CH230-49F
C 30	51	41.1	749	17	AZ234421	AZ234421 RPI1-23-9
C 31	51	41.1	946	17	CNS03MJG	AL250837 Tetraodon
C 32	51	41.1	1022	17	CNS02HD6	AL197475 Tetraodon
C 33	51	41.1	1148	12	BG330598	AL937475 Tetraodon
C 34	50.5	40.7	300	9	AJ477360	AJ477360 AJ477360
C 35	50.5	40.7	500	17	BH871063	BH871063 hm70b08.g
C 36	50	40.3	304	10	AW007889	AW007889 wv46e09.x
C 37	50	40.3	374	14	BQ101082	BQ101082 lj23g05.y
C 38	50	40.3	389	10	AW275500	AW275500 xn07h11.x
C 39	50	40.3	398	10	AW304845	AW304845 xv96b07.x
C 40	50	40.3	423	14	BM678760	BM678760 UI-E-EJ0
C 41	50	40.3	432	9	AI034279	AI034279 ox89e10.x
C 42	50	40.3	436	14	W66232	W66232 TgESTy71b1
C 43	50	40.3	459	14	BQ185659	BQ185659 UI-E-EJ1
C 44	50	40.3	477	17	AZ895789	AZ895789 RPI1-24-2
C 45	50	40.3	487	10	AW024849	AW024849 wu69g07.x

ALIGNMENTS

RESULT 1  
BF911801/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BF911801  
MR4-UT0051-101100-001-a10 UT0051 Homo sapiens CDNA, mRNA sequence.  
BF911801  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 443)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

443 bp mRNA linear EST 18-JAN-2001

TITLE	Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000.)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&t2=MR4-UT0051- 101100-001-al08t3-2008-11-10&t4=1) Seq primer: puc 18 forward High quality sequence start: 21 High quality sequence stop: 443.

FEATURES	SOURCE
1. <b>High Accuracy:</b> The model achieves a high accuracy rate, indicating its effectiveness in predicting the target variable.	Model Performance Metrics
2. <b>Robustness:</b> The model demonstrates robustness against various input variations and noise, ensuring reliable predictions.	Model Performance Metrics
3. <b>Interpretability:</b> The model's predictions are interpretable, allowing users to understand the underlying factors influencing the outcome.	Model Performance Metrics
4. <b>Scalability:</b> The model is scalable, capable of handling large volumes of data and complex tasks efficiently.	Model Performance Metrics
5. <b>Real-time Processing:</b> The model supports real-time data processing, enabling immediate decision-making based on the latest information.	Model Performance Metrics
6. <b>Integration with Existing Systems:</b> The model seamlessly integrates with existing data management and analytics systems.	Model Performance Metrics
7. <b>Customizable Parameters:</b> The model offers customizable parameters to tailor its performance to specific use cases.	Model Performance Metrics
8. <b>Regular Updates:</b> The model is regularly updated with the latest data and algorithms to maintain its predictive accuracy.	Model Performance Metrics
9. <b>Compliance with Regulations:</b> The model adheres to relevant data protection and privacy regulations, ensuring ethical data usage.	Model Performance Metrics
10. <b>Support for Multiple Data Sources:</b> The model can ingest data from multiple sources, providing a comprehensive view of the data landscape.	Model Performance Metrics

```

source
I. .443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0051"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 136,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      112 a      86 c      113 g      132 t
ORIGIN

Alignment Scores:
Pred. No.:      9.39      Length:      443
Score:      58.00      Matches:      12
Percent Similarity:      80.95%      Conservative:      5
Best Local Similarity:      57.14%      Mismatches:      3
Query Match:      46.77%      Indels:      1
DB:      12      Gaps:      0

US-09-980-523A-6 (1-22) x BF911801 (1-443)

```

Qy	3	GluValTyrAspPhe-AlaPheArgAspLeuCysIleValTyrArgAspGlyAsnProty	22
		:::           :::           :::           :::	
Db	320	GAGGTATTGAATTGCGCCITTAAGATTTATTGTGTGTATAGACACAGTATACCGCA	261
Qy	22	r	22
Db	260	T	260
RESULT 2			
BG810071/c			
LOCUS	BG810071	681 bp	linear
DEFINITION	mgct002xxk19f Magnaporthe grisea Appressorium stage cDNA Magnaporthe grisea cDNA clone mgct002xxk19f 5', mRNA sequence.		
ACCESSION	BG810071		
VERSION	BG810071.1	GI:14181051	
KEYWORDS	EST.		
SOURCE	Magnaporthe grisea.		
ORGANISM	Magnaporthe grisea		
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.		
AUTHORS	1 (bases 1 to 681)		
TITLE	Choi,W. and Dean,R.A.		
JOURNAL	Construction and sequence analysis of an appressorium stage cDNA library in the rice blast fungus, Magnaporthe grisea		
	Unpublished (2001)		

COMMENT

Contact: Ralph A. Dean  
Fungal Genomics Laboratory  
North Carolina State University  
Campus Box 7251, Raleigh, NC 27695, USA  
Tel: 919-513-0020  
Tel: 919-513-0020  
Fax: 919-513-0024  
Email: [ralph\\_dean@ncsu.edu](mailto:ralph_dean@ncsu.edu)  
Seq primer: T3 primer (AATTAAACCTCCTCACTAAAGGG)

Alignment Scores:		
Pred. No.:	15	9
Score:	58.00	
Percent Similarity:	68.18%	
Best Local Similarity:	45.45%	
Query Match:	46.77%	
DB:	12	
Length:		68
Matches:		10
Conservative:		5
Mismatches:		7
Indels:		0
Gaps:		0

RESULT 3  
AV862231/C

FEATURES	SOURCE
----------	--------



ACCESSION AV852103  
 VERSION AV852103.1  
 KEYWORDS GI:16836951  
 SOURCE EST.  
 ORGANISM Ciona intestinalis.

REFERENCE 1 (bases 1 to 520)  
 AUTHORS Sato, N., Satou, Y., Kohara, Y. and Shin-i.T.  
 TITLE Expressed genes in Ciona intestinalis  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Nori Sato  
 Department of Zoology  
 Kyoto University  
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4081  
 Fax: 81-75-705-1113  
 Email: sato@cica.kyoto-u.ac.jp

US-09-980-523A-6 (1-22) x AV862231 (1-520)

QY 3 GluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsnProTyR 22  
 DB 323 GAAGTCGGATCCTTCATCTCCGCTCATCTGCATCATCTATAGTGTGGGGAGGCAC 264

#### RESULT 4

AV868962/c  
 LOCUS AV868962  
 DEFINITION Nori Sato unpublished cDNA library, egg Ciona intestinalis cDNA clone rcieg35g03 3', mRNA sequence.

ACCESSION AV868962  
 VERSION AV868962.1  
 KEYWORDS GI:16856486  
 SOURCE EST.

ORGANISM Ciona intestinalis.  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 Phlebobranchia; Clonidae; Ciona.

REFERENCE 1 (bases 1 to 586)  
 AUTHORS Sato, N., Satou, Y., Kohara, Y. and Shin-i.T.  
 TITLE Expressed genes in Ciona intestinalis  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Nori Sato  
 Department of Zoology  
 Kyoto University  
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4081  
 Fax: 81-75-705-1113  
 Email: sato@cica.kyoto-u.ac.jp

#### FEATURES

source  
 1..586  
 /organism="Ciona intestinalis"  
 /db\_xref="taxon:7719"  
 /clone="rcieg35g03"  
 /tissue\_type="whole animal"  
 /dev\_stage="egg"  
 /note="Vector: pBluescript SK"

BASE COUNT 121 a 141 c 144 g 180 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 28.3 Length: 586  
 Score: 56.00 Matches: 10  
 Percent Similarity: 65.00% Conservative: 3  
 Best Local Similarity: 50.00% Mismatches: 7  
 Query Match: 45.16% Indels: 0  
 DB: 10 Gaps: 0

US-09-980-523A-6 (1-22) x AV868962 (1-586)

QY 3 GluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsnProTyR 22  
 DB 328 GAAGTCGGATCCTTCATCTCCGCTCATCTGCATCATCTATAGTGTGGGGAGGCAC 269

#### RESULT 5

AV852103/c  
 LOCUS AV852103  
 DEFINITION Nori Sato unpublished cDNA library, egg Ciona intestinalis cDNA clone rcieg20g03 3', mRNA sequence.

ACCESSION AV852103  
 VERSION AV852103.1  
 KEYWORDS GI:16836951  
 SOURCE EST.  
 ORGANISM Ciona intestinalis.

REFERENCE 1 (bases 1 to 622)  
 AUTHORS Sato, N., Satou, Y., Kohara, Y. and Shin-i.T.  
 TITLE Expressed genes in Ciona intestinalis  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Nori Sato  
 Department of Zoology  
 Kyoto University  
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4081  
 Fax: 81-75-705-1113  
 Email: sato@cica.kyoto-u.ac.jp

FEATURES  
 source  
 1..622  
 /organism="Ciona intestinalis"  
 /db\_xref="taxon:7719"  
 /clone="rcieg20g03"  
 /tissue\_type="whole animal"  
 /dev\_stage="egg"  
 /note="Vector: pBluescript SK"

BASE COUNT 134 a 146 c 150 g 184 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 30.4 Length: 622  
 Score: 56.00 Matches: 10  
 Percent Similarity: 65.00% Conservative: 3  
 Best Local Similarity: 50.00% Mismatches: 7  
 Query Match: 45.16% Indels: 0  
 DB: 10 Gaps: 0

US-09-980-523A-6 (1-22) x AV852103 (1-622)

QY 3 GluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsnProTyR 22  
 DB 306 GAAGTCGGATCCTTCATCTCCGCTCATCTGCATCATCTATAGTGTGGGGAGGCAC 247

RESULT 6  
 AG166938/c  
 LOCUS AG166938  
 DEFINITION Pan troglodytes DNA, clone: RP43-034N17.TJ, genomic survey sequence.

ACCESSION AG166938  
 VERSION AG166938.1  
 KEYWORDS GI:16696616  
 SOURCE GSS.

ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1  
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 TITLE BAC end sequences of Library RPCI-43  
 JOURNAL Unpublished  
 COMMENT 2 (bases 1 to 758)  
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 Direct Submission  
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan 305-8565, Japan  
 (E-mail: chimpesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library RPCI-43 this BAC end was generated during the R&D process and may have higher chance

of clone tracking errors.

## PRIMERS

Sequencing: TJ

## LIBRARY

Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.

## FEATURES

source

Location/Qualifiers

1..758

/organism="Pan troglodytes"

/db\_xref="taxon:9598"

/clone="RP43-034N17.TJ"

/sex="male"

/cell\_type="lymphocytes"

/clone\_lib="RPGI-43 Chimpanzee Male BAC Library"

250 a 148 c 112 g 247 t 1 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 56.7 Length: 758  
Score: 55.00 Matches: 11  
Percent Similarity: 63.64% Conservatives: 3  
Best Local Similarity: 50.00% Mismatches: 6  
Query Match: 44.35% Indels: 2  
DB: 17 Gaps: 1

US-09-980-523A-6 (1-22) x AG166938 (1-758)

QY 1 ArgArgGluValTyArgPheAlaPheArgLeuCysIleValTyArgGspGlyAsn 20

Db 177 AGGAGAGAATCTACTTCATCTTTT-----CTTGGGTAGTATTCATAGTCTAAAT 124

QY 21 ProTyr 22

Db 123 CCATAT 118

RESULT 7

B20885/c

LOCUS

DEFINITION B20885 825 bp DNA linear GSS 16-SEP-1997  
F11E8-T7.2 IGF Arabidopsis thaliana genomic clone F11E8, DNA

ACCESSION

VERSION B20885.1 GI:2395939

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 825)

Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and

Ecker,J.

BAC End Sequences at ATGC

Unpublished (1997)

Other\_GSSs: F11E8-T7, F11E8-Sp6, F11E8-T7.1

Contact: Ecker J.

Arabidopsis Thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA

19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 156

High quality sequence stop: 191.

Location/Qualifiers

1..825

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db\_xref="taxon:3702"

/clone="F11E8"

## FEATURES

source

Alignment Scores:

Pred. No.: 69 Length: 890  
Score: 55.00 Matches: 11  
Percent Similarity: 80.95% Conservatives: 6  
Best Local Similarity: 52.38% Mismatches: 3  
Query Match: 44.35% Indels: 1

/clone\_lib="IGF"

/sex="hermaphrodite"

/note="Vector: BelOBACII; Site\_1: EcoRI; Site\_2: EcoRI;

Produced by Thomas Altmann"

260 a 170 c 103 g 264 t 28 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 62.9 Length: 825  
Score: 55.00 Matches: 10  
Percent Similarity: 68.18% Conservatives: 5  
Best Local Similarity: 45.43% Mismatches: 7  
Query Match: 44.35% Indels: 0  
DB: 17 Gaps: 0

US-09-980-523A-6 (1-22) x B20885 (1-825)

QY 1 ArgArgGluValTyArgPheAlaPheArgLeuCysIleValTyArgGspGlyAsn 20

Db 489 CGCCGAGAGATGNCAGAGACTTTTCTTGAGAGATATACCCATTATTTTGAGATGNCAC 430

QY 21 ProTyr 22

Db 429 CCATTC 424

RESULT 8

BES43005

LOCUS

DEFINITION

890 bp mRNA linear EST 09-AUG-2000

601068927F1 NIH\_MGC\_12 Homo sapiens CDNA clone IMAGE:3455214 5',

mrna sequence.

ACCESSION BES43005

VERSION BES43005.1 GI:9771650

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 890)

NIH-MGC http://mgi.nci.nih.gov/.

Unpublished (1999)

National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ARCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM8441 row: e column: 07

High quality sequence stop: 542.

Location/Qualifiers

1..890

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3455214"

/clone\_lib="NIH\_MGC\_12"

/tissue\_type="cervical carcinoma cell line"

/lab\_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.4 kb. Library prepared by Life

Technologies."

339 a 220 c 174 g 157 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 69 Length: 890  
Score: 55.00 Matches: 11  
Percent Similarity: 80.95% Conservatives: 6  
Best Local Similarity: 52.38% Mismatches: 3  
Query Match: 44.35% Indels: 1





```
BE067679
LOCUS      BE067679          644 bp      mRNA      linear      EST 09-JUN-2000
DEFINITION MR4-BT0358-270300-005-f09 BT0358 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE067679
VERSION    BE067679.1 GI:8412329
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 644)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
COMMENT    Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MR4-BT0358-270
300-005-f09&t3=2000-03-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 101.
FEATURES   Location/Qualifiers
            source          1..644
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_lib="BT0358"
                        /dev_stage="Adult"
                        /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                        SmaI; A mini-library was made by cloning products derived
                        from ORESTES PCR (U.S. Letters Patent application No. 196
                        716 - Ludwig Institute for Cancer Research) profiles
                        into the pUC 18 vector. Reverse transcription of tissue
                        mRNA and cDNA amplification were performed under low
                        stringency conditions."
BASE COUNT 214 a 135 c 108 g 187 t
ORIGIN
Alignment Scores:
Pred. No.: 99.7 Length: 644
Score: 53.00 Matches: 10
Percent Similarity: 55.56% Conservative: 0
Best Local Similarity: 55.56% Mismatches: 8
Query Match: 42.74% Indels: 0
DB: 10 Gaps: 0
US-09-980-523a-6 (1-22) x BE067679 (1-644)
QY 5 TyraSpPheAlaPheArgAspLeuCySileValTyArgAspGlyAsnProTyTr 22
Db 408 TATTCTGTCTCTGAGAGGGCTCTGCACCTGTATACAGGTCAGGATACCATAC 461
RESULT 14
AG029226
LOCUS      AG029226          755 bp      DNA      linear      GSS 01-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-001D23.R, genomic survey sequence.
ACCESSION  AG029226
VERSION    AG029226.1 GI:16556098
KEYWORDS   GSS.
Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-001D23.R.
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 755)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpesegsc.riken.go.jp, URL:http://hgsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES   Location/Qualifiers
            source          1..755
                        /organism="Pan troglodytes"
                        /db_xref="taxon:9598"
                        /clone="PTB-001D23.R"
                        /sex="male"
                        /cell_type="lymphoblast"
                        /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 191 a 125 c 134 g 303 t
ORIGIN
Alignment Scores:
Pred. No.: 121 Length: 755
Score: 53.00 Matches: 8
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 57.14% Mismatches: 3
Query Match: 42.74% Indels: 0
DB: 17 Gaps: 0
US-09-980-523a-6 (1-22) x AG029226 (1-755)
QY 4 ValTyAspPheAlaPheArgAspLeuCySileValTyArg 17
Db 148 GTTATGTCATTGTCATTGAAGGACATTGTGTTGTTATAGG 189
RESULT 15
BE539653/c
LOCUS      BE539653/c          1044 bp      mRNA      linear      EST 09-AUG-2000
DEFINITION 601059444F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3446051 5',
mRNA sequence.
ACCESSION  BE539653
VERSION    BE539653.1 GI:9768298
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1044)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
CONTACT: Robert Strausberg, Ph.D.
Email: cgapus@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 6, 2003, 12:41:15 ; Search time 217 Seconds  
(without alignments)  
228.313 Million cell updates/sec

Title: US-09-980-523A-6

Perfect score: 124

Sequence: 1 RREYDFAFRLCIVYRDGPY 22

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-O=/Cgn2\_1/USP20\_spool/US09980523/runat\_03072003\_164405\_4208/app\_query.fasta\_1.199  
-DB=N.Geneseq\_101002 -QFMT=tastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -NATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTENT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09980523 -RCGN\_1\_1\_396 @runat\_03072003\_164405\_4208 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	519	17	AAT31835 Human papilloma vi
2	124	100.0	570	16	AAQ75470 HPV16 E6/E7 encodi
3	124	100.0	712	10	AA91600 Partial nucleotide
4	124	100.0	776	17	AAT14663 E6/E7 region of Hu
5	124	100.0	779	20	AA89756 Probe sequence for
6	124	100.0	790	13	AAQ29389 DNA encoding HPV 1
7	124	100.0	801	17	AAT31833 Human papilloma vi
8	124	100.0	822	20	AA87872 HPV fusion protein
9	124	100.0	822	20	AA87872 Prot.D1/3-E6-His/H
10	124	100.0	879	20	AA87875 HPV fusion protein
11	124	100.0	879	20	AA87875 CLYTA-E6-His codin
12	124	100.0	921	21	AA09477 Human papillomavir
13	124	100.0	939	24	AA09477 Human papillomavir
14	124	100.0	1005	10	AA91784 DNA probe compleme
15	124	100.0	1116	20	AA87873 HPV fusion protein
16	124	100.0	1116	20	AA87873 Prot.D1/3-E6-E7-Hi
17	124	100.0	1173	20	AA87877 HPV fusion protein
18	124	100.0	1173	20	AA87877 CLYTA-E6E7-His cod
19	124	100.0	7840	22	AAF55127 Nucleotide sequenc
20	124	100.0	7902	17	AAT09847 Human papilloma vi
21	124	100.0	7904	18	AAT94723 Human papillomavir
22	124	100.0	7904	18	AAT94724 Human papillomavir
23	124	100.0	7904	20	AA83388 HPV-16 genomic seq
24	124	100.0	7904	20	AA83388 Nucleotide sequenc
25	124	100.0	7904	20	AA83388 Nucleotide sequenc
26	117	94.4	1000	11	AA08627 HPV-16 fragment.
27	94	75.8	1095	20	AA833781 HPV68 DNA sequence
28	94	75.8	1108	20	AA833781 HPV68 E6 and E7 pr
29	93	75.0	774	10	AA91602 Partial nucleotide
30	93	75.0	1023	14	AAQ48575 HPV E6/E7 Genomic r
31	89	71.8	3191	11	AAQ40092 3.2 kb KpnI-PvuII
32	86	66.1	3283	20	AAZ31789 HPV70 DNA sequence
33	79	63.7	7833	13	AAZ31789 Human papilloma vi
34	71	57.3	476	16	AAQ97848 HPV-18 E6 cDNA. H
35	71	57.3	476	20	AAZ27568 HPV-18 E6 coding s
36	71	57.3	476	21	AAA61623 HPV18 E6/E7 protei
37	71	57.3	483	16	AAQ75471 DNA encoding HPV-
38	71	57.3	817	13	AAQ29390 HPV18 E6/E7 protei
39	71	57.3	837	20	AA87880 HPV fusion protein
40	71	57.3	837	20	AA87880 Prot.D1/3-E6-His/H
41	71	57.3	1000	11	AAQ08623 HPV-18 fragment.
42	71	57.3	1152	20	AA87880 HPV fusion protein
43	71	57.3	1152	20	AA87880 Prot.D1/3-E6-E7-Hi
44	71	57.3	1731	9	AA80161 Human papilloma vi
45	65	52.4	50	10	AA891799 DNA oligonucleotid

# ALIGNMENTS

RESULT 1

AAT31835

ID AAT31835 standard; DNA; 519 BP.

XX

AC AAT31835;

XX

DT 11-JAN-1997 (first entry)

XX

DE Human papilloma virus E6/E7 protein variant.

XX

KW Human papilloma virus; E6; E7; deletion mutant; HPV;

KW immune response; humoral immune response; cellular immune response;

KW vaccine; ss.

XX

OS Human papilloma virus.

XX

XX

FT Key Location/Qualifiers

FT CDS 1..519

FT /\*tag= a  
FT /product= E6/E7 fusion protein.  
FT  
PN W09619496-A1.  
XX  
XX 27-JUN-1996.  
PD  
XX  
XX 20-DEC-1995; 95WO-AU00868.  
XX  
XX 20-DEC-1994; 94AU-0000157.  
PR  
XX  
XX (CSLC-) CSL LTD.  
PA (UYQU ) UNIV QUEENSLAND.  
PA  
XX  
XX Cox J, Edwards SJ, Frazer I, Webb EA;  
PI  
XX WPI; 1996-309518/31.  
DR P-PSDB; AAR97563.  
XX

XX Vaccine variants of human papilloma virus antigens - contain  
PT variants of E6 and/or E7 protein, pref. deletion mutants, and are  
PT used to treat or prevent HPV infection  
XX

PS Example 3; Page 18; 37pp; English.

XX A variant of the human papilloma virus (HPV) E6 or E7 protein which  
CC elicits a humoral and/or cellular immune response against HPV can be  
CC used in vaccines against HPV or to treat HPV infection. The variant  
CC is preferably a deletion mutant comprising at least half, and  
CC preferably two-thirds of full length E6 or E7 protein starting from  
CC the N- or C-terminal, or is a full length E6 moiety fused to a full  
CC length E7 moiety. The variant optionally has a linkage moiety and a  
CC foreign protein or peptide which facilitates the purification of,  
CC and enhances the immunogenicity of, the fusion protein. This  
CC sequence encodes a fusion between the C-terminal end of E7 and the  
CC N-terminal end of E6.

XX Sequence 519 BP; 167 A; 97 C; 111 G; 144 T; 0 other;

Alignment Scores:  
Pred. No.: 6.26e-13 Length: 519  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-09-980-523A-6 (1-22) x AAT31835 (1-519)

Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
Db 343 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGATGGGAAT 402

Qy 21 ProTyr 22

Db 403 CCATAT 408

RESULT 2

ID AAQ75470  
XX AAQ75470 standard; DNA; 570 BP.

XX AC AAQ75470;

XX 28-JUN-1995 (first entry)

XX HPV16 E6/E7 encoding region.

XX HPV; HPV16; E6 protein; E7 protein; diagnosis; cervical dysplasia;  
XX cervix cancer; ds.

XX Human papillomavirus strain 16.

XX Key Location/Qualifiers

CDS 2..478  
FT /\*tag= .a  
FT /label= E6\_encoding\_region  
FT 481..570  
FT /\*tag= b  
FT /label= E7\_encoding\_region

XX W09426934-A.

XX 24-NOV-1994.

XX 06-MAY-1994; 94WO-US05085.

XX 06-MAY-1993; 93US-0058920.

XX (BAXT ) BAXTER DIAGNOSTICS INC.

XX Brown JT;

XX WPI; 1995-006821/01.

XX P-PSDB; AAR63865.

XX Human papilloma virus detection assay - by amplification using  
PT self sustained sequence replication and hybridisation with a  
PT detector probe

XX Disclosure; Page 24-26; 79pp; English.

XX The sequences of the E6 and E7 polypeptide-encoding regions of human  
CC papillomavirus (HPV) 16 and 18 are given in AAQ75470-71 and the  
CC encoded proteins in AAR63865-66, respectively. Probes and primers  
CC based on these sequences were used for HPV infection diagnosis;  
CC expression of E6 and E7 is diagnostic for cervical cancer or pre-  
CC malignant states.

XX Sequence 570 BP; 199 A; 93 C; 119 G; 159 T; 0 other;

Alignment Scores:  
Pred. No.: 7.02e-13 Length: 570  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0

US-09-980-523A-6 (1-22) x AAQ75470 (1-570)

Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
Db 137 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGATGGGAAT 196

Qy 21 ProTyr 22

Db 197 CCATAT 202

RESULT 3

ID AAN91600  
XX AAN91600 standard; DNA; 712 BP.

XX AC AAN91600;

XX 17-JUL-1990 (first entry)

XX Partial nucleotide sequence (5' end) of human papilloma virus (HPV)  
DE type 16 (HPV-16).

XX Human papilloma virus; type 16; in situ hybridisation assay;  
XX cellular smear; benign cervical wart; cervical cancer.

XX Human papilloma virus.

XX W08902934-A.



```

PD 06-APR-1989.
XX
XX 30-SEP-1988; 88WO-US03367.
XX
XX 02-OCT-1987; 87US-0103979.
XX
XX (MCR-) MICROPROBE CORP.
XX
XX Schwartz DE, Adams TH;
XX
XX WPI; 1989-114406/15.
XX
XX Hybridisation test for human papilloma virus in cell smears -
XX by reaction with long labelled probe specific for particular
XX virus types, esp. for examining cervical smears
XX
XX Disclosure; ; 39pp; English.
XX
XX The patent is for a rapid in situ hybridisation assay for detecting and
XX typing human papilloma virus (HPV) in non-frozen cellular smears fixed to
XX a support in absence of aldehyde-based crosslinking reagents. The assay
XX comprises: (1) combining nucleic acid in the sample with at least one
XX detectable probe able to hybridise with 1 or more HPV types; and (2)
XX detecting presence or absence of hybrid complexes. Opt. several probes
XX are used, eg one for HPV types 6 and 11, associated with benign warts,
XX and one for types 16, 18, 31, 33 and 35, associated with cervical
XX cancer. The assay can differentiate between HPV types. It is esp. used
XX as a secondary test. The probes can be synthesised or cloned.
XX
XX Sequence 712 BP; 253 A; 120 C; 155 G; 184 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 9.2e-13 Length: 712
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-980-523A-6 (1-22) x AAN91600 (1-712)
Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCystIleValTyrArgAspGlyAsn 20
Db 218 CGACGTGAGGTATGACTTGTCTTTCGGCAITTTATGCATATATAGATGGGAAT 277
Qy 21 ProTyr 22
Db 278 CCATAT 283

RESULT 4
AAT14663
ID AAT14663 standard; DNA; 776 BP.
XX
XX AC AAT14663;
XX
XX 10-OCT-1996 (first entry)
XX
XX E6/E7 region of Human Papilloma Virus 16 (HPV 16).
XX
XX Human papilloma virus; HPV; detection; cervical cancer;
XX amplification; hybridisation; diagnosis; transformed cell; E6; E7;
XX
XX Human papilloma virus 16.
XX
XX Key Location/Qualifiers
XX primer_bind 1..24
XX /tag= a
XX /note= "Primer BB113 binding site."
XX misc_binding 30..55
XX /tag= b
XX /note= "Primer H16-58 binding site."
XX primer_bind 37..57

```

```

FT /tag= c
FT /note= "Primer BB4 binding site."
FT primer_bind 454..474
FT /tag= d
FT /note= "Primer BB114 binding site."
FT primer_bind 480..503
FT /tag= e
FT /note= "Primer BB111 binding site."
FT primer_bind 591..621
FT /tag= f
FT /note= "Primer BB109 binding site."
FT primer_bind 658..681
FT /tag= g
FT /note= "Primer H16-686 binding site."
FT primer_bind 660..683
FT /tag= h
FT /note= "Primer BB112 binding site."
FT primer_bind 715..738
FT /tag= i
FT /note= "Primer H16-743 binding site."
FT primer_bind 745..768
FT /tag= j
FT /note= "Primer H16-773 binding site."
XX
XX US5506105-A.
XX
XX 09-APR-1996.
XX
XX 10-DEC-1991; 91US-0808456.
XX
XX 10-DEC-1991; 91US-0808456.
XX 22-MAR-1994; 94US-0216233.
XX (DADE-) DADE INT INC.
XX Haydock PV;
XX WPI; 1996-200273/20.
XX
XX Detection of low copy number intracellular markers - by 3SR
XX amplification of target RNA in fixed cells then hybridisation with
XX labelled probe, for detecting human papilloma virus in cervical
XX cells
XX
XX Example 2; Figure 3; 21pp; English.
XX
XX An in situ hybridisation assay for detecting an intracellular marker
XX of low copy number in cells comprises fixing the cells to a support
XX using paraformaldehyde; treating the cells with a protease to
XX permeabilise them without altering morphology; adding amplification
XX reagents; incubating the cells at below fifty degrees celcius to
XX perform amplification by self-sustained sequence replication; adding
XX a labelled probe complementary to the region between the primers;
XX washing cells to remove unhybridised probe and then detecting the
XX labelled probe. The E6/E7 region of human papilloma virus (HPV) 16
XX is used especially to detect mRNA being transcribed from this
XX region which becomes active in transformed cells. The method can be
XX used for the early diagnosis of cervical cancer. Primers used to
XX amplify fragments of the E6/E7 region are described in
XX AAT14664-T14674.
XX
XX Sequence 776 BP; 260 A; 139 C; 169 G; 208 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 1.02e-12 Length: 776
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-980-523A-6 (1-22) x AAT14663 (1-776)

```

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 136 CGACGTGAGGTATGACTTTCGTTTCGGGATTATGCATAGTATAGATGGGAAT 195  
QY 21 ProTyr 22  
|||||  
Db 196 CCATAT 201

## RESULT 5

AAX89756

ID AAX89756 standard; DNA; 779 BP.

XX AC

XX AAX89756;

XX AC

XX 12-OCT-1999 (first entry)

XX DE

XX Probe sequence for HPV 16 E6/E7.

XX KW human papilloma virus; infection; gene expression; probe;

XX KW detection; assay; cancer; virus; HPV; ss.

XX XX

XX Synthetic.

XX OS Human Papilloma virus.

XX PN

XX WO9929890-A2.

XX PD

XX 17-JUN-1999.

XX PF

XX 11-DEC-1998; 98WO-US26447.

XX PR

XX 17-APR-1998; 98US-0082167.

XX PR

XX 12-DEC-1997; 97US-0069426.

XX PR

XX 05-JAN-1998; 98US-0070486.

XX PA (DIGE-) DIGENE CORP.

XX XX

XX Lorincz AT;

XX WPI; 1999-443850/37.

XX DR

XX New method for assessing Human Papilloma Virus (HPV) infection by

XX comparison of gene expression levels

XX PS Disclosure; Fig 5; 35pp; English.

XX XX

XX This nucleotide probe is specific for the HPV16 Human Papilloma

XX Virus (HPV) gene.

XX CC The degree of HPV infection can be assessed, by measuring

XX CC the levels of expression of genes involved in the diseased state, and

XX CC comparing the expression to each other or to reference genes.

XX CC

XX Sequence 779 BP; 259 A; 139 C; 170 G; 211 T; 0 other;

XX XX

## Alignment Scores:

Pred. No.: 1.03e-12 Length: 779

Score: 124.00 Matches: 22

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 20 Gaps: 0

US-09-980-523A-6 (1-22) x AAX89756 (1-779)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20

|||||

Db 143 CGACGTGAGGTATGACTTTCGTTTCGGGATTATGCATAGTATAGATGGGAAT 202

QY 21 ProTyr 22

|||||

Db 203 CCATAT 208

RESULT 6

AAX29389

AAQ29389 standard; DNA; 790 BP.

XX AC

XX AAQ29389;

XX 09-MAR-1993 (first entry)

XX DT

XX DNA encoding HPV 16 E6/E7 proteins obtd. by PCR.

XX DE

XX Virus vector; vaccinia virus; papillomavirus; HPV; human;

XX KW amplification; immunotherapeutic; ss.

XX KW

XX Human papillomavirus 16.

XX OS

XX Key

XX Location/Qualifiers

XX FT 1..789

XX CDS /tag= a

XX /note= "reading frame 1"

XX FT 2..790

XX CDS /tag= b

XX /note= "reading frame 2"

XX FT 485..781

XX misc\_feature /tag= c

XX /note= "second reading frame encoding HPV 16 E7"

XX FT 3..788

XX CDS /tag= d

XX /note= "reading frame 3"

XX FT 6..482

XX misc\_feature /tag= e

XX /note= "third reading frame encoding HPV 16 E6"

XX FT

XX WO9216636-A..

XX PN

XX 01-OCT-1992.

XX PD

XX 10-MAR-1992; 92WO-GB00424.

XX PF

XX 14-MAR-1991; 91GB-0005383.

XX PR

XX (IMMU ) IMMUNOLOGY LTD.

XX PA

XX Bournsneil MEG, Inglis SC, Munro AJ;

XX PI

XX WPI; 1992-349219/42.

XX DR

XX P-PSDB; AAR27723-5.

XX DR

XX Recombinant virus vectors encoding human papillomavirus proteins

XX - for treating and vaccinating against HPV infections and

XX conditions caused by them, such as cervical cancer

XX PS Disclosure; Fig 1a; 83pp; English.

XX XX

XX The fragment of DNA contg. the HPV-16 E6/E7 coding region was

XX prepd. by PCR from plasmid pBR322/HPV16 (Durst et al., PNAS, 80;

XX 3612 (1983)) using oligonucleotides S05 and S06. The DNA prod. has

XX a site for NcoI at the beginning of the E6 gene and a SmaI site

XX immediately downstream of the termination codon for E7. The E6 and

XX E7 ORFs are fused together to form a single continuous ORF via site

XX directed mutagenesis and the immortalising potential of E7 is removed

XX by altering two key codons of the HPV E7 sequence. The single ORF of

XX HPV-16 E6/E7 may be inserted into vaccinia virus DNA at neutral sites

XX (pref. by inserting two sets of the DNA in opposite orientations to

XX overcome the problem of intertypic recombination) to make a recombinant

XX virus vector for use immunotherapeutically to activate cells of the

XX immune system against HPV. See also AAQ29385-400 and AAQ29450-69.

XX SQ Sequence 790 BP; 262 A; 144 C; 175 G; 209 T; 0 other;

XX XX

XX Alignment Scores:

Pred. No.: 1.04e-12 Length: 790

Score: 124.00 Matches: 22

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

```
DB: 13 Gaps: 0
US-09-980-523A-6 (1-22) x AAQ29389 (1-790)
Qy 1 ArgArgGluValTyArgAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20
Db 141 CGAGCTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATATATAGATAGATGGGAAT 200
Qy 21 ProTyr 22
Db 201 CCATAT 206
RESULT 7
AAT31833
ID AAT31833 standard; DNA; 801 BP.
AC AAT31833;
XX
DT 10-JAN-1997 (first entry)
DE Human papilloma virus E6/E7 protein variant.
XX
KW Human papilloma virus; E6; E7; deletion mutant; HPV;
KW immune response; humoral immune response; cellular immune response;
KW vaccine; ss.
XX
OS Human papilloma virus.
FH
FI Key Location/Qualifiers
FT CDS 1..801
FT FT /*tag= a
FT FT /product= E6/E7 fusion protein.
XX
PN W09619496-A1.
XX
PD 27-JUN-1996.
XX
PF 20-DEC-1995; 95WO-AU00868.
XX
PR 20-DEC-1994; 94AU-0000157.
XX
PA (CSLC-) CSL LTD.
PA (UYQU ) UNIV QUEENSLAND.
XX
PI Cox J, Edwards SJ, Frazer I, Webb EA;
XX
DR WPI; 1996-309518/31.
DR P-PSDB; AAR97561.
XX
XX Vaccine variants of human papilloma virus antigens - contain
PT variants of E6 and/or E7 protein, pref. deletion mutants, and are
PT used to treat or prevent HPV infection
XX
PS Example 1; Page 15-16; 37pp; English.
XX
CC A variant of the human papilloma virus (HPV) E6 or E7 protein which
CC elicits a humoral and/or cellular immune response against HPV can be
CC used in vaccines against HPV or to treat HPV infection. The variant
CC is preferably a deletion mutant comprising at least half, and
CC preferably two-thirds of full length E6 or E7 protein starting from
CC the N- or C-terminal, or is a full length E6 moiety fused to a full
CC length E7 moiety. The variant optionally has a linkage moiety and a
CC foreign protein or peptide which facilitates the purification of,
CC and enhances the immunogenicity of, the fusion protein. This
CC sequence encodes a full length E6/E7 fusion protein.
XX
SQ Sequence 801 BP; 266 A; 150 C; 174 G; 211 T; 0 other;
Alignment Scores:
Pred No.: 1,06e-12 Length: 801
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-09-980-523A-6 (1-22) x AAT31833 (1-801)
Qy 1 ArgArgGluValTyArgAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20
Db 136 CGAGCTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATATATAGATAGATGGGAAT 195
Qy 21 ProTyr 22
Db 196 CCATAT 201
RESULT 8
AAX78792
ID AAX78792 standard; DNA; 822 BP.
XX
AC AAX78792;
DT 06-SEP-1999 (first entry)
DE HPV fusion protein D1/3-E6-His/HPV16 DNA.
XX
KW Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW immunological fusion partner; CpG oligonucleotide; immune response;
KW HPV antigen; prevention; treatment; ss.
XX
OS Synthetic.
OS Human papillomavirus.
XX
PN W09933868-A2.
PD 08-JUL-1999.
XX
PF 18-DEC-1998; 98WO-EP08563.
XX
PR 24-DEC-1997; 97GB-0027262.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Dalemans WLJ, Gerard CMG;
XX
DR WPI; 1999-405485/34.
DR P-PSDB; AAY25376.
XX
PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT induce immune response to HPV
XX
PS Example II; Page 48; 62pp; English.
XX
CC AAX78791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY25375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory CpG
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used
CC for preventing or treating HPV induced tumours.
XX
SQ Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 other;
Alignment Scores:
Pred No.: 1.1e-12 Length: 822
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0
US-09-980-523A-6 (1-22) x AAX78792 (1-822)
Qy 1 ArgArgGluValTyArgAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20
Db 454 CGAGCTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATATATAGATAGATGGGAAT 513
```

```
QY      21 ProtYr 22
AC      |||||
DB      514 CCATAT 519

RESULT 9
AAAX29781
ID AAX29781 standard; DNA; 822 BP.
XX
AC AAX29781;
XX
DT 22-JUN-1999 (first entry)
XX
DE Prot.D1/3-E6-His/HPV16 coding sequence.
XX
KW Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
KW tumour; lesion; benign; malignant; virus; infection; ss.
XX
XX Chimeric - Human papillomavirus.
OS Chimeric - Haemophilus influenzae.
XX
PN WO9910375-A2.
XX
PD 04-MAR-1999.
XX
PF 17-AUG-1998; 98WO-EP05285.
XX
PR 22-AUG-1997; 97GB-0017953.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck C, Cabazon Silva T, Delisse AEF, Gerard CMG;
PI Lombardo-Bencheikh A;
XX
DR WPI; 1999-190587/16.
XX
XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
PT treatment or prophylaxis of HPV induced lesions
PT
XX Disclosure; Fig 3; 95pp; English.
XX
XX This sequence represents the coding region for a chimeric E6 or E7
CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked
CC to an immunological fusion partner, in this case, a fragment of the
CC Haemophilus influenzae B protein D. The sequence also contains a
CC histidine tag at the C-terminus of the encoded protein. The protein
CC can be used in a vaccine, for immuno-therapeutically treating HPV
CC induced tumour lesions (benign or malignant) and preventing HPV viral
CC infection.
XX
SQ Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 other;

Alignment Scores:
Pred. No.: 1.1e-12 Length: 822
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-980-523A-6 (1-22) x AAX29781 (1-822)

QY      1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
DB      454 CGACGTGAGGTATGACTTTCGTTTCGGGATTTATGCATAGTATATAGATGGGAAT 513
XX
QY      21 ProtYr 22
AC      |||||
DB      514 CCATAT 519

RESULT 10
AAAX78795
ID AAX78795 standard; DNA; 879 BP.
XX
```

```
XX AAX78795;
AC
DT 06-SEP-1999 (first entry)
XX
DE HPV fusion protein CLYTA-E6-His/HPV16 DNA.
XX
KW Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW immunological fusion partner; CpG oligonucleotide; immune response;
KW HPV antigen; prevention; treatment; ss.
XX
OS Synthetic.
OS Human papillomavirus.
XX
PN WO9933868-A2.
XX
PD 08-JUL-1999.
XX
PF 18-DEC-1998; 98WO-EP08563.
XX
PR 24-DEC-1997; 97GB-0027262.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Dalemans WLJ, Gerard CMG;
XX
DR WPI; 1999-405485/34.
DR P-PSDB; AAY25379.
XX
PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT induce immune response to HPV
XX
PS Example VI; Page 52; 62pp; English.
XX
XX AAX78791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY25375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory CpG
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used
CC for preventing or treating HPV induced tumours.
XX
SQ Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 other;

Alignment Scores:
Pred. No.: 1.19e-12 Length: 879
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-980-523A-6 (1-22) x AAX78795 (1-879)

QY      1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
DB      511 CGACGTGAGGTATGACTTTCGTTTCGGGATTTATGCATAGTATATAGATGGGAAT 570
XX
QY      21 ProtYr 22
AC      |||||
DB      571 CCATAT 576

RESULT 11
AAAX29784
ID AAX29784 standard; DNA; 879 BP.
XX
AC AAX29784;
XX
DT 22-JUN-1999 (first entry)
XX
DE CLYTA-E6-His coding sequence.
XX
KW Chimeric; E6; E7; fusion protein; CLYTA; vaccine; immunotherapy;
```



```

DE Human papillomavirus enhE6,7 construct DNA.
XX
KW Alphavirus vector system: human papilloma virus; HPV; cervical cancer;
KW therapy; vaccine; virucide; cytostatic; ds.
XX
OS Human papillomavirus.
XX
PN EP1195438-A1.
XX
PD 10-APR-2002.
XX
XX 06-OCT-2000; 2000EP-0203472.
XX
XX 06-OCT-2000; 2000EP-0203472.
XX
XX (UYGR-) RIJKSUNIV GRONINGEN.
XX
XX Regts DG, Wilschut JC, Holtrop M, Daemen CAHH;
PI WPI; 2002-354156/39.
XX
XX New alphavirus system, useful for genetic immunization against cervical
XX cancer, comprises papilloma virus nucleic acid -
XX
XX Example 2; Fig 19; 45pp; English.
XX
XX The present invention relates to an alphavirus vector system comprising
XX nucleic acid derived from a human papilloma virus (HPV). The invention
XX or cells containing it, are used in treatment and prevention of cervical
XX cancer, particularly as a vaccine. By selecting the nucleic acid that
XX encode E6/E7 proteins without ability to bind to pRB and p53, the risk
XX that cells infected with the alphavirus vector system may become
XX oncogenic is avoided (contrast use of other viral vectors). The present
XX sequence is Human papillomavirus enhE6,7 construct DNA.
XX
XX Sequence 939 BP; 284 A; 197 C; 217 G; 241 T; 0 other;
SQ

```

Alignment Scores:

Pred. No.:	1.29e-12	Length:	939
Score:	124.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-980-523A-6 (1-22) x AAN91784 (1-939)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysTleValTyrArgAspGlyAsn 20  
 |||||  
 DB 300 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTATGTCATATATAGATGGGAT 359  
 |||||  
 QY 21 ProTyr 22  
 |||||  
 DB 360 CCATAT 365

RESULT 14

AAN91784  
 ID AAN91784 standard; DNA: 1005 BP.  
 AC AAN91784;  
 XX  
 XX 16-MAR-1990 (first entry)  
 DT  
 XX DNA probe complementary to human papilloma virus type 16.  
 DE  
 XX Cervical cancer.  
 KW  
 XX Human papilloma virus type 16.  
 OS  
 XX W08909940-A.  
 PN  
 XX 19-OCT-1989.  
 PD  
 XX

US-09-980-523A-6 (1-22) x AAN91784 (1-1005)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
 |||||  
 DB 670 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTATGTCATATATAGATGGGAT 729  
 |||||  
 QY 21 ProTyr 22  
 |||||  
 DB 730 CCATAT 735

RESULT 15

AAAX78793  
 ID AAAX78793 standard; DNA: 1116 BP.  
 XX  
 AC AAAX78793;  
 XX  
 XX 06-SEP-1999 (first entry)  
 DT  
 XX HPV fusion protein DL/3-E6E7-His/HPV16 DNA.  
 DE  
 XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;  
 KW immunological fusion partner; CpG oligonucleotide; immune response;  
 KW HPV antigen; prevention; treatment; ss.  
 XX  
 XX Synthetic.  
 OS  
 XX Human papillomavirus.  
 OS  
 XX W09933868-A2.  
 PN  
 XX 08-JUL-1999.  
 PD  
 XX 18-DEC-1998; 98WO-EP08563.  
 PF  
 XX 24-DEC-1997; 97GB-0027262.  
 PR  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 XX Dalemans WLJ, Gerard CMG;  
 PI  
 XX WPI; 1999-405485/34..  
 DR P-PSDB; AAY25377.  
 DR  
 XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to  
 XX induce immune response to HPV  
 PT

PF 04-APR-1989; 89WO-US01318.  
 XX  
 PR 04-APR-1988; 88US-0177404.  
 XX  
 PA (ONCO-) ONCOR INC.  
 XX  
 XX George AL, Groff DE;  
 PI WPI; 1989-324314/44.  
 DR  
 XX Rapid detection of specific human papilloma virus genotypes - by  
 PT hybridisation of DNA digest with new labelled nucleic acid probes  
 PT  
 XX Claim 40; page 46; 81pp; English.  
 PS  
 XX Obtd. by cutting HPV16 with BamHI and PvuII. The patent describes probes  
 CC (DNA or RNA) and their complements capable of detecting one or a  
 CC combination of HPV types 6, 11, 16, 18, 31, 33 and 35.  
 CC  
 XX Sequence 1005 BP; 325 A; 182 C; 190 G; 308 T; 0 other;  
 SQ

Alignment Scores:

Pred. No.:	1.4e-12	Length:	1005
Score:	124.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

```
XX Example III; Page 49; 62pp; English.
PS
XX AAX78791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY25375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory CpG
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used
CC for preventing or treating HPV induced tumours.
XX
SQ Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 other;

Alignment Scores:
Pred. No.:      1.59e-12      Length:      1116
Score:          124.00      Matches:      22
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              20      Gaps:      0

US-09-980-523A-6 (1-22) x AAX78793 (1-1116)
QY      1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db      454 CGACGTGAGGTATATGACTTTGCTTTTCGGATTATGCATAGTATATAGAGATGGGAAT 513
QY      21 ProTyr 22
Db      514 CCATAT 519
```

Search completed: July 6, 2003, 13:12:03  
Job time : 219 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 16:39:57 ; Search time 26 Seconds  
(without alignments)  
24.896 Million cell updates/sec

Title: US-09-980-523A-6

Perfect score: 124

Sequence: 1 RREYDFAFRDLCIVYRDGNPY 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*\*

1: /cgn2.6/prodata/1/1aa/5A\_COMB.pep.\*

2: /cgn2.6/prodata/1/1aa/5B\_COMB.pep.\*

3: /cgn2.6/prodata/1/1aa/6A\_COMB.pep.\*

4: /cgn2.6/prodata/1/1aa/6B\_COMB.pep.\*

5: /cgn2.6/prodata/1/1aa/PTCUS\_COMB.pep.\*

6: /cgn2.6/prodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	162	1 US-08-316-239B-3	Sequence 3, Appl
2	124	100.0	162	1 US-08-316-239B-4	Sequence 4, Appl
3	124	100.0	172	3 US-08-860-165-14	Sequence 14, Appl
4	124	100.0	172	4 US-09-359-382-14	Sequence 14, Appl
5	124	100.0	266	3 US-08-860-165-10	Sequence 10, Appl
6	124	100.0	266	4 US-09-359-382-10	Sequence 10, Appl
7	124	100.0	266	4 US-09-367-309A-1	Sequence 1, Appl
8	124	100.0	273	4 US-09-485-885-4	Sequence 4, Appl
9	124	100.0	292	4 US-09-485-885-10	Sequence 10, Appl
10	124	100.0	371	4 US-09-485-885-6	Sequence 6, Appl
11	124	100.0	390	4 US-09-485-885-14	Sequence 14, Appl
12	97	78.2	20	2 US-08-934-915-161	Sequence 161, App
13	80	64.5	182	1 US-08-117-083-10	Sequence 10, Appl
14	72	58.1	14	1 US-07-509-122-4	Sequence 4, Appl
15	71	57.3	158	2 US-08-247-904B-10	Sequence 10, Appl
16	71	57.3	158	3 US-08-767-942A-19	Sequence 19, Appl
17	71	57.3	271	1 US-08-117-083-14	Sequence 14, Appl
18	71	57.3	278	4 US-09-485-885-21	Sequence 21, Appl
19	71	57.3	383	4 US-09-485-885-23	Sequence 23, Appl
20	56	45.2	368	4 US-09-000-094-20	Sequence 20, Appl
21	56	45.2	375	4 US-09-000-094-22	Sequence 22, Appl
22	56	45.2	465	4 US-09-000-094-24	Sequence 24, Appl
23	56	45.2	1587	4 US-08-000-094-46	Sequence 46, Appl
24	52	41.9	9	3 US-08-159-339A-133	Sequence 133, App
25	50	40.3	1162	4 US-09-134-001C-4008	Sequence 4008, Ap
26	48	38.7	9	1 US-08-787-547-103	Sequence 103, App
27	48	38.7	9	3 US-08-159-339A-246	Sequence 246, App

28	48	38.7	9	3	US-08-159-339A-564	Sequence 564, App
29	48	38.7	9	4	US-08-197-484-67	Sequence 67, Appl
30	48	38.7	9	5	PCT-US95-02121-67	Sequence 67, Appl
31	48	38.7	484	1	US-08-111-939-12	Sequence 12, Appl
32	48	38.7	719	4	US-09-641-741-28	Sequence 28, Appl
33	48	38.7	845	4	US-09-641-741-29	Sequence 29, Appl
34	48	38.7	1128	1	US-08-111-939-2	Sequence 2, Appl
35	48	38.7	1128	4	US-09-641-741-30	Sequence 30, Appl
36	48	38.7	1128	4	US-09-060-482-8	Sequence 8, Appl
37	48	38.7	1158	4	US-09-060-482-2	Sequence 2, Appl
38	45	36.3	263	4	US-08-776-059-43	Sequence 43, Appl
39	45	36.3	264	4	US-08-776-059-33	Sequence 33, Appl
40	45	36.3	564	4	US-08-159-339A-219	Sequence 219, App
41	44	35.5	9	3	US-09-122-443-16	Sequence 16, Appl
42	42	33.9	185	3	US-09-230-637-25	Sequence 25, Appl
43	42	33.9	204	4	US-09-230-371A-26	Sequence 26, Appl
44	42	33.9	204	4	US-09-230-371A-26	Sequence 26, Appl
45	42	33.9	710	4	US-09-079-812E-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-316-239B-3  
Sequence 3, Application US/08316239B  
Patent No. 5679509  
GENERAL INFORMATION:  
APPLICANT: Wheeler, Cosette M.  
APPLICANT: Parmenter, Cheryl A.  
TITLE OF INVENTION: Methods and a Diagnostic Aid for  
Distinguishing a Subset of HPV that is Associated with an  
Increased Risk of Developing Cervical Dysplasia and  
Cervical Cancer  
TITLE OF INVENTION: Cervical Cancer  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jagtiani & Associates  
STREET: 6126 Rocky Way Court  
CITY: Centreville  
STATE: VA  
COUNTRY: USA  
ZIP: 20120-3400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,239B  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jagtiani, Ajay A.  
REGISTRATION NUMBER: 35,205  
REFERENCE/DOCKET NUMBER: UNME-0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 817-9453  
TELEFAX: (703) 803-9387  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 162 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-316-239B-3

Query Match 100.0%; Score 124; DB 1; Length 162;  
Best Local Similarity 100.0%; Pred. No. 5.4e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22

```

; EARLIER APPLICATION NUMBER: AU PN0157
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-14

Query Match 100.0%; Score 124; DB 3; Length 172;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 4
US-09-359-382-14
; Sequence 14, Application US/09359382
; Patent No. 6306397
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 017227/0148
; CURRENT APPLICATION NUMBER: US/09/359,382
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 08/860,165
; EARLIER FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157/94
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-359-382-14

Query Match 100.0%; Score 124; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 5
US-08-860-165-10
; Sequence 10, Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 17227/130
; CURRENT APPLICATION NUMBER: US/08/860,165A
; CURRENT FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157

```

```

; EARLIER APPLICATION NUMBER: AU PN0157
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-14

Query Match 100.0%; Score 124; DB 3; Length 172;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 4
US-09-359-382-14
; Sequence 14, Application US/09359382
; Patent No. 6306397
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 017227/0148
; CURRENT APPLICATION NUMBER: US/09/359,382
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 08/860,165
; EARLIER FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157/94
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-359-382-14

Query Match 100.0%; Score 124; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 5
US-08-860-165-10
; Sequence 10, Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 17227/130
; CURRENT APPLICATION NUMBER: US/08/860,165A
; CURRENT FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157

```

```

; EARLIER APPLICATION NUMBER: AU PN0157
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-14

Query Match 100.0%; Score 124; DB 3; Length 172;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 4
US-09-359-382-14
; Sequence 14, Application US/09359382
; Patent No. 6306397
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 017227/0148
; CURRENT APPLICATION NUMBER: US/09/359,382
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 08/860,165
; EARLIER FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157/94
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-359-382-14

Query Match 100.0%; Score 124; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 5
US-08-860-165-10
; Sequence 10, Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 17227/130
; CURRENT APPLICATION NUMBER: US/08/860,165A
; CURRENT FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157

```

US-08-860-165-10  
OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion

; SOFTWARE: PatentIn Ver. 2.1

```
; ORGANISM: Homo sapien
US-09-485-885-10

Query Match      100.0%; Score 124; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 171 RREYDFAFRDLCIVYRDGNPY 192

RESULT 10
US-09-485-885-6
; Sequence 6, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-6

Query Match      100.0%; Score 124; DB 4; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 152 RREYDFAFRDLCIVYRDGNPY 173

RESULT 11
US-09-485-885-14
; Sequence 14, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-14

Query Match      100.0%; Score 124; DB 4; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
Db 171 RREYDFAFRDLCIVYRDGNPY 192

RESULT 12
US-08-934-915-161
; Sequence 161, Application US/08934915
; Patent No. 5932412
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; APPLICANT: CHENG, HWEI-MING
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES
; NUMBER OF SEQUENCES: 193
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MASON & ASSOCIATES, P.A.
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
; CITY: CLEARWATER
; STATE: FLORIDA
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 3.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,915
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/949,836
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LOUISE A. FOUTCH
; REGISTRATION NUMBER: 37,133
; REFERENCE/DOCKET NUMBER: 1946.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 813-538-3800
; TELEFAX: 813-538-3820
; TELEX:
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-934-915-161

Query Match      78.2%; Score 97; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 1.2e-09;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 REYDFAFRDLCIVYRDGNP 21
Db 1 REYDFAFRDLCIVYRDGNP 20

RESULT 13
US-08-117-083-10
; Sequence 10, Application US/08117083
; Patent No. 5719054
; GENERAL INFORMATION:
; APPLICANT: Boursnell, Michael E.
; APPLICANT: Inglis, Stephen C.
```

APPLICANT: Munro, Alan J.  
TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human  
Papilloma Virus Proteins  
NUMBER OF SEQUENCES: 70  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Walter H. Dreger  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/117,083  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Walter H.  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-58783  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 182 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..182  
OTHER INFORMATION: /note= "xaa refers to stop codon in  
the open reading frame."  
US-08-117-083-10

Query Match 64.5%; Score 80; DB 1; Length 182;  
Best Local Similarity 84.2%; Pred. No. 1.1e-05;  
Matches 16; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 RREYDFAFRDLCLV--YR 17  
Db 47 RREYDFAFRDLCLTISEYR 65

RESULT 14  
US-07-909-122-4  
Sequence 4, Application US/07909122  
Patent No. 5415995  
GENERAL INFORMATION:  
APPLICANT: SCHOOLNIK, GARY K.  
APPLICANT: PALEFSKY, JOEL M.  
TITLE OF INVENTION: DIAGNOSTIC PEPTIDES OF HUMAN PAPILLOMA  
VIRUS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/909,122  
FILING DATE: 19920706  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BENZ, WILLIAM H.  
REGISTRATION NUMBER: 25,952  
REFERENCE/DOCKET NUMBER: 28600-20105.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-909-122-4

Query Match 58.1%; Score 72; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 9.9e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCL 13  
Db 2 RREYDFAFRDLCL 14

RESULT 15  
US-08-247-904B-10  
Sequence 10, Application US/08247904B  
Patent No. 5981699  
GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Eckstein, Jens W.  
APPLICANT: Draetta, Giulio  
TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley, Hoag & Elliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,904B  
FILING DATE: 23-MAY-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-247-904B-10

Query Match 57.3%; Score 71; DB 2; Length 158;  
Best Local Similarity 60.0%; Pred. No. 0.00027;  
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
QY      3  EYDFAFRDLCIVYRDGNY  22
        |::|::|:|::|:|:|
Db      43  EVFEFAFKDLFVVYRDSIPH  62
```

Search completed: July 3, 2003, 16:44:29  
Job time : 27 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	124	100.0	273	9	US-10-000-903-4		Sequence 4, Appli
2	124	100.0	292	9	US-10-000-903-10		Sequence 10, Appl
3	124	100.0	371	9	US-10-000-903-6		Sequence 6, Appli
4	124	100.0	390	9	US-10-000-903-14		Sequence 14, Appl
5	71	57.3	278	9	US-10-000-903-21		Sequence 21, Appl
6	71	57.3	383	9	US-10-000-903-23		Sequence 23, Appl
7	48	38.7	9	9	US-09-909-460-103		Sequence 103, App
8	48	38.7	9	9	US-10-138-711-67		Sequence 67, Appl
9	48	38.7	9	9	US-10-133-210-281		Sequence 861, App
10	48	38.7	153	10	US-09-925-301-856		Sequence 285, App
11	48	38.7	719	9	US-09-996-015-9		Sequence 9, Appli
12	48	38.7	845	9	US-09-996-015-11		Sequence 11, Appl
13	48	38.7	1128	9	US-09-996-015-10		Sequence 10, Appl
14	48	38.7	1128	9	US-10-238-876-8		Sequence 8, Appli
15	48	38.7	1158	9	US-10-238-876-2		Sequence 2, Appli
16	47	37.9	164	9	US-10-136-761-11918		Sequence 11918, A
17	47	37.9	554	10	US-09-746-491-41		Sequence 41, Appl
18	46	37.1	216	10	US-09-811-284-181		Sequence 181, App
19	46	37.1	653	10	US-09-801-574-82		Sequence 82, Appl

; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/10/000,903  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 292  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-000-903-10

Query Match 100.0%; Score 124; DB 9; Length 292;  
Best Local Similarity 100.0%; Pred. No. 6.9e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
Db 171 RREYDFAFRDLCIVYRDGNPY 192

## RESULT 3

US-10-000-903-6  
; Sequence 6, Application US/10000903  
; Publication No. US20020182221A1

; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/10/000,903  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-000-903-6

Query Match 100.0%; Score 124; DB 9; Length 371;  
Best Local Similarity 100.0%; Pred. No. 9e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
Db 152 RREYDFAFRDLCIVYRDGNPY 173

## RESULT 4

US-10-000-903-14  
; Sequence 14, Application US/10000903  
; Publication No. US20020182221A1

; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela

; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/10/000,903  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 390  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-000-903-14

Query Match 100.0%; Score 124; DB 9; Length 390;  
Best Local Similarity 100.0%; Pred. No. 9.5e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
Db 171 RREYDFAFRDLCIVYRDGNPY 192

## RESULT 5

US-10-000-903-21  
; Sequence 21, Application US/10000903  
; Publication No. US20020182221A1

; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/10/000,903  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21  
; LENGTH: 278  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-000-903-21

Query Match 57.3%; Score 71; DB 9; Length 278;  
Best Local Similarity 60.0%; Pred. No. 0.0018;  
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EYDFAFRDLCIVYRDGNPY 22  
Db 154 EVFEFAFKDLFVYVYRDSIPH 173

## RESULT 6

US-10-000-903-23  
; Sequence 23, Application US/10000903  
; Publication No. US20020182221A1

; GENERAL INFORMATION:  
; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/10/000,903



;; CURRENT FILING DATE: 2001-10-01  
;; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
;; PRIOR FILING DATE: 1998-08-17  
;; PRIOR APPLICATION NUMBER: GB 9717953.5  
;; PRIOR FILING DATE: 1997-08-22  
;; NUMBER OF SEQ ID NOS: 23  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 23  
;; LENGTH: 383  
;; TYPE: PRT  
;; ORGANISM: Homo sapien  
US-10-000-903-23

Query Match 57.3%; Score 71; DB 9; Length 383;  
Best Local Similarity 60.0%; Pred. No. 0.0025;  
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EVYDFAFDLCIVYRDCNPY 22  
||:||||:||||:|  
Db 154 EVFEAFKDLFVYRDSIPH 173

## RESULT 7

US-09-909-460-103  
;; Sequence 103, Application US/09909460  
;; Publication No. US20020182258A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lunsford, Lynn B.  
;; APPLICANT: Hedley, Mary Lynn  
;; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF NUCLEIC  
;; TITLE OF INVENTION: ACID  
;; FILE REFERENCE: 08191/014001  
;; CURRENT APPLICATION NUMBER: US/09/909,460  
;; CURRENT FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/321,346  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
;; NUMBER OF SEQ ID NOS: 114  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 103  
;; LENGTH: 9  
;; TYPE: PRT  
;; ORGANISM: Human papilloma virus  
US-09-909-460-103

Query Match 38.7%; Score 48; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.9e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FAFRDLCIV 15  
|||||  
Db 1 FAFRDLCIV 9

## RESULT 8

US-10-128-711-67  
;; Sequence 67, Application US/10128711  
;; Publication No. US20030099634A1  
;; GENERAL INFORMATION:  
;; APPLICANT: VITIELLO, Maria A.  
;; APPLICANT: CHESTNUT, Robert W.  
;; APPLICANT: SETTE, Alessandro D.  
;; APPLICANT: CELIS, Esteban  
;; APPLICANT: GRAY, Howard  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
;; TITLE OF INVENTION: CTL IMMUNITY  
;; NUMBER OF SEQUENCES: 153  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend Kourie and Crew  
;; STREET: Stewart Street Tower, One Market Plaza  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: US

ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/128,711  
FILING DATE: 22-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/197,484  
FILING DATE: 16-FEB-1994  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (206) 623-6793  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 67:

US-10-128-711-67  
Query Match 38.7%; Score 48; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.9e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FAFRDLCIV 15  
|||||  
Db 1 FAFRDLCIV 9

## RESULT 9

US-10-133-210-281  
;; Sequence 281, Application US/10133210  
;; Publication No. US20030103964A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Delisi, Charles  
;; APPLICANT: Berzofsky, Jay  
;; APPLICANT: Gulukota, Kamalakara  
;; APPLICANT: Vaccaro, Dennis  
;; APPLICANT: Weng, Zhiping  
;; APPLICANT: Zhang, Chao  
;; TITLE OF INVENTION: METHODS FOR DESIGNING MOLECULAR CONJUGATES AND  
;; TITLE OF INVENTION: COMPOSITIONS THEREOF  
;; FILE REFERENCE: BU-035AX  
;; CURRENT APPLICATION NUMBER: US/10/133,210  
;; CURRENT FILING DATE: 2002-04-26  
;; NUMBER OF SEQ ID NOS: 281  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 281  
;; LENGTH: 9  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-133-210-281

Query Match 38.7%; Score 48; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.9e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FAFRDLCLIV 15  
|||||  
Db 1 FAFRDLCLIV 9

RESULT 10  
US-09-925-301-856  
; Sequence 856, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 856  
; LENGTH: 153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-856

Query Match 38.7%; Score 48; DB 10; Length 153;  
Best Local Similarity 50.0%; Pred. No. 4.2;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 REVYDFAFRDLCLIVRDGNP 21  
||: :|| |||||  
Db 118 REILLMLQYLCREYRDGNP 137

RESULT 11  
US-09-996-015-9  
; Sequence 9, Application US/09996015  
; Publication No. US20030032166A1  
; GENERAL INFORMATION:  
; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Pena, Carol A. E.  
; APPLICANT: Li, Li  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Leite, Marlo W.  
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids  
; FILE REFERENCE: 15966-581 CIP  
; CURRENT APPLICATION NUMBER: US/09/996,015  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: 09/641,741  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 60/159,613  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: 60/175,534  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 60/224,086  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 719  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-996-015-9

Query Match 38.7%; Score 48; DB 9; Length 719;  
Best Local Similarity 50.0%; Pred. No. 23;

Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
QY 2 REVYDFAFRDLCLIVRDGNP 21  
||: :|| |||||  
Db 214 REILLMLQYLCREYRDGNP 233

RESULT 12  
US-09-996-015-11  
; Sequence 11, Application US/09996015  
; Publication No. US20030032166A1  
; GENERAL INFORMATION:  
; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Pena, Carol A. E.  
; APPLICANT: Li, Li  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Leite, Marlo W.  
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids  
; FILE REFERENCE: 15966-581 CIP  
; CURRENT APPLICATION NUMBER: US/09/996,015  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: 09/641,741  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 60/159,613  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: 60/175,534  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 60/224,086  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 845  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-996-015-11

Query Match 38.7%; Score 48; DB 9; Length 845;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 REVYDFAFRDLCLIVRDGNP 21  
||: :|| |||||  
Db 319 REILLMLQYLCREYRDGNP 338

RESULT 13  
US-09-996-015-10  
; Sequence 10, Application US/09996015  
; Publication No. US20030032166A1  
; GENERAL INFORMATION:  
; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Pena, Carol A. E.  
; APPLICANT: Li, Li  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Leite, Marlo W.  
; TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids  
; FILE REFERENCE: 15966-581 CIP  
; CURRENT APPLICATION NUMBER: US/09/996,015  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: 09/641,741  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: 60/159,613  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: 60/175,534  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 60/224,086  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10

; LENGTH: 1128  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-996-015-10

Query Match 38.7%; Score 48; DB 9; Length 1128;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 REVYDFAFRDLCIVYRDGNP 21  
||: :|| |||||  
Db 623 RELLLLMQYLCQYRDGNP 642

## RESULT 14

US-10-238-876-8  
; Sequence 8, Application US/10238876  
; Publication No. US20030084464A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Mu-En  
; APPLICANT: Layne, Matthew D.  
; APPLICANT: Yet, Shaw-Pang  
; TITLE OF INVENTION: AORTIC CARDOXYPEPTIDASE-LIKE POLYPEPTIDE  
; FILE REFERENCE: 21509-036CIPCON  
; CURRENT APPLICATION NUMBER: US/10/238,876  
; CURRENT FILING DATE: 2002-09-10  
; PRIOR APPLICATION NUMBER: US 09/060,482  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: US 08/818,009  
; PRIOR FILING DATE: 1997-03-14  
; PRIOR APPLICATION NUMBER: US 60/013,439  
; PRIOR FILING DATE: 1996-03-15  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 1128  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-238-876-8

Query Match 38.7%; Score 48; DB 9; Length 1128;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 REVYDFAFRDLCIVYRDGNP 21  
||: :|| |||||  
Db 623 RELLLLMQYLCQYRDGNP 642

## RESULT 15

US-10-238-876-2  
; Sequence 2, Application US/10238876  
; Publication No. US20030084464A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Mu-En  
; APPLICANT: Layne, Matthew D.  
; APPLICANT: Yet, Shaw-Pang  
; TITLE OF INVENTION: AORTIC CARDOXYPEPTIDASE-LIKE POLYPEPTIDE  
; FILE REFERENCE: 21509-036CIPCON  
; CURRENT APPLICATION NUMBER: US/10/238,876  
; CURRENT FILING DATE: 2002-09-10  
; PRIOR APPLICATION NUMBER: US 09/060,482  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: US 08/818,009  
; PRIOR FILING DATE: 1997-03-14  
; PRIOR APPLICATION NUMBER: US 60/013,439  
; PRIOR FILING DATE: 1996-03-15  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1158  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-238-876-2

Query Match 38.7%; Score 48; DB 9; Length 1158;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 REVYDFAFRDLCIVYRDGNP 21  
||: :|| |||||  
Db 632 RELLLLMQYLCQYRDGNP 651

Search completed: July 3, 2003, 16:52:41  
Job time : 52 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 16:39:56 ; Search time 70 Seconds  
(without alignments)  
41.879 Million cell updates/sec

Title: US-09-980-523A-6

Perfect score: 124

Sequence: 1 RREYDFAFRLCIVIRGNPY 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	22	22	Peptide comprising
2	124	100.0	23	20	Amino acid sequenc
3	124	100.0	23	20	HPV-derived lipope
4	124	100.0	22	22	HPV 16 E6 protein
5	124	100.0	32	22	Polyepitopic fragm
6	124	100.0	32	22	A polyepitopic fra
7	124	100.0	117	22	Polyepitopic polype
8	124	100.0	151	21	HPV-16 E6 protein
9	124	100.0	158	21	HPV E6 peptide. S
10	124	100.0	158	21	Human papillomavir

11	124	100.0	158	22	AA98420	Human papillomavir
12	124	100.0	162	18	AA35741	Human papillomavir
13	124	100.0	162	18	AA35742	Human papillomavir
14	124	100.0	172	17	AA97563	Human papilloma vi
15	124	100.0	188	16	AA63865	HPV16 E6/E7 protei
16	124	100.0	236	22	AA96052	Polyepitope polype
17	124	100.0	243	20	AA99369	Papillomavirus E6/
18	124	100.0	263	13	AA27725	HPV 16 E6 protein
19	124	100.0	266	17	AA97561	Human papilloma vi
20	124	100.0	273	20	AA25376	HPV fusion protein
21	124	100.0	273	20	AA25379	Prot.D1/3-E6-His/H
22	124	100.0	292	20	AA25379	HPV fusion protein
23	124	100.0	292	20	AA25379	CLYTA-E6-His prote
24	124	100.0	371	20	AA25377	HPV fusion protein
25	124	100.0	371	20	AA25377	Prot.D1/3-E6-E7-Hi
26	124	100.0	390	20	AA25381	HPV fusion protein
27	124	100.0	390	20	AA25381	CLYTA-E6E7-His pro
28	98	79.0	19	19	AA29371	Peptide Gf56. Syn
29	93	75.0	149	14	AA40919	HPV E6 region prod
30	93	75.0	149	22	AA98447	Human papillomavir
31	72	58.1	14	9	AA81958	Peptide immunogen
32	72	58.1	149	22	AA98434	Human papillomavir
33	72	58.1	158	22	AA98440	Human papillomavir
34	71	57.3	158	16	AA79656	HPV-18 E6 protein.
35	71	57.3	158	16	AA63866	HPV18 E6/E7 protei
36	71	57.3	158	20	AA39968	HPV-18 E6 protein
37	71	57.3	158	21	AA803176	HPV-18 E6 protein.
38	71	57.3	158	22	AA98427	Human papillomavir
39	71	57.3	271	13	AA27728	HPV 18 E6 protein
40	71	57.3	278	20	AA25385	HPV fusion protein
41	71	57.3	278	20	AA25386	Prot.D1/3-E6-His/H
42	71	57.3	383	20	AA25386	HPV fusion protein
43	71	57.3	383	20	AA25386	Prot.D1/3-E6-E7-Hi
44	65	52.4	155	22	AA98452	Human papillomavir
45	63	50.8	16	19	AA29370	Peptide Gf55. Syn

#### ALIGNMENTS

##### RESULT 1

AA30943

ID AA30943 standard; peptide; 22 AA.

XX AA30943;

XX AA30943;

DT 02-APR-2001 (first entry)

XX Peptide comprising amino acids 46-67 of E6 protein of HPV.

DE Polypeptidic peptide; E6 protein; E7 protein; HPV; CD4 epitope;

XX T helper cell; human leukocyte antigen; HLA; immune response; cytotoxic T cell; CTL; cytokine secretion; interleukin-2; IL-2; IL-4;

KW gamma-interferon; HPV infection; cervical neoplasia; invasive cancer;

KW vulvar intraepithelial neoplasia.

XX Human papilloma virus.

OS FR2794371-A1.

XX 08-DEC-2000.

PD 07-OCT-1999; 99FR-0012511.

XX 03-JUN-1999; 99FR-0007012.

XX (BIOV-) BIOVECTOR THERAPEUTICS SA.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Choppin J, Bourgault VI, Guillet JG, Connan F, Ferries E;

XX WPI; 2001-064175/08.

DR

XX

PT New polypeptidic fragments from human papilloma virus E6 and E7  
 PT proteins, useful for treatment or prevention of e.g. cervical neoplasia  
 PT and cancer -

PS Claim 2; Page 21; 27pp; French.

XX AAB30942-80 represent polypeptidic fragments from the E6 and E7 proteins  
 CC of human papilloma virus (HPV). The peptides include CD4 epitopes  
 CC recognised by T helper cells. They bind stably to human leukocyte  
 CC antigen (HLA) type molecules. The peptides induce a specific immune  
 CC response, particularly cytotoxicity, caused by cytotoxic T cells (CTL).  
 CC They also induce secretion of cytokines (particularly interleukin-2  
 CC (IL-2) and IL-4, and gamma-interferon) by CTL. The peptides, their  
 CC derivatives, nucleic acids encoding them and specific antibodies are  
 CC used, in compositions or vaccines, to treat or prevent diseases  
 CC associated with HPV infection, e.g. cervical or vulvar intraepithelial  
 CC neoplasia and invasive cancer of the cervix uteri. The antibodies are  
 CC also useful for in vitro diagnosis of these diseases.

XX Sequence 22 AA;

Query Match 100.0%; Score 124; DB 22; Length 22;

Best Local Similarity 100.0%; Pred. No. 7.2e-14;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22

Db 1 RREYDFAFRDLCIVYRDGNPY 22

RESULT-2

AY40374

ID AAY40374 standard; Peptide; 23 AA.

AC AAY40374;

XX 19-NOV-1999 (first entry)

XX Amino acid sequence of an E6 and E7 protein epitope.

XX Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;  
 KW CD4+ epitope; B epitope; lipopeptide; interferon gamma; adjuvant;  
 KW vaccine; tumor; infection; immune response; cytokine profile;  
 KW acquired immune deficiency syndrome; papilloma; cancer; hepatitis;  
 KW autoimmune disease.

XX Human papillomavirus.

OS FR2774687-A1.

PN 13-AUG-1999.

PD 06-FEB-1998; 98FR-0001439.

PF 06-FEB-1998; 98FR-0001439.

PR (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

PA (INSP ) INST PASTEUR LILLE.

XX Thiam K, Guillet JG, Ver Waerde C, Auriault C, Gras MH, Loing E;

XX WPI; 1999-510734/43.

XX New lipopeptide comprising C-terminal interferon-gamma fragment with  
 PT attached lipophilic groups, used as interferon mimic, e.g. for treating  
 PT cancer or virus infection

PS Disclosure; Page 42; 53pp; French.

XX AAY40123-Y40379 represent epitopes that are able to activate cytotoxic  
 CC T lymphocytes (CD8+ epitopes), T helper cells (CD4+ epitopes), or  
 CC B epitopes recognized by corresponding antibodies. The epitopes may be  
 CC used in the composition of the invention. The specification describes a

CC lipopeptide that has a peptide part derived from mammalian interferon  
 CC gamma (IFN $\gamma$ ) and one or more lipophilic parts comprising a linear or  
 CC branched, (un)saturated 4-20C hydrocarbon chain or a steroid. The  
 CC lipopeptide mimics the activity of IFN $\gamma$ . Compositions comprising the  
 CC lipopeptide are used to treat or prevent any condition that responds  
 CC to IFN $\gamma$ , and as adjuvant for vaccines (particularly those directed  
 CC against tumors, viral or parasitic infections), to stimulate or  
 CC (re)orient the immune response between types 1 and 2 cytokine profiles.  
 CC Particular applications are treatment of infections (particularly  
 CC viral, e.g. acquired immune deficiency syndrome, papilloma (cancer) and  
 CC hepatitis, but also bacterial, fungal, parasitic or helminth); cancers  
 CC (particularly of kidney, cutaneous T cells or ovary, chronic  
 CC myelogenous leukemia or mesothelioma), allergy; and autoimmune  
 CC diseases.

XX Sequence 23 AA;

Query Match 100.0%; Score 124; DB 20; Length 23;

Best Local Similarity 100.0%; Pred. No. 7.5e-14;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22

Db 2 RREYDFAFRDLCIVYRDGNPY 23

RESULT 3

AAY26725

ID AAY26725 standard; peptide; 23 AA.

XX AAY26725;

AC AAY26725;

XX 14-SEP-1999 (first entry)

XX HPV-derived lipopeptide epitope (E6 aa45-67) for mixed micelles.

XX Micelle; microaggregate; induction; immune response; lipopeptide; CTL;  
 KW cytotoxic T-lymphocyte; epitope; lipid; helper T-lymphocyte; HTL; HBV;  
 KW tetanus; toxin; vaccine; HIV; hepatitis B virus; papilloma virus; p53;  
 KW melanoma; Plasmodium falciparum; malaria.

XX Synthetic.

OS Human papillomavirus.

XX FR2771640-A1.

PN 04-JUN-1999.

XX 03-DEC-1997; 97FR-0015246.

PF 03-DEC-1997; 97FR-0015246.

PR (CNRS ) CNRS CENT NAT RECH SCI.

PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

PA (INSP ) INST PASTEUR LILLE.

XX Bossus M, Bourgault VL, Gras-Masse H, Guillet JG, Lippens G;

PI Tartar A, Wieruszkeski JM;

XX WPI; 1999-349509/30.

XX Immunogenic lipopeptide micelles - comprising lipopeptides  
 PT containing cytotoxic and helper T-lymphocyte epitopes

PS Disclosure; Page 32; 60pp; French.

XX The invention relates to the generation of mixed micelles or  
 CC microaggregates for inducing an immune response comprise: (a) a first  
 CC lipopeptide comprising at least one CTL (cytotoxic T-lymphocyte) epitope  
 CC and at least one lipid unit; and (b) a second lipopeptide comprising at  
 CC least one HTL (helper T-lymphocyte) epitope and at least one lipid unit  
 CC different from that of the first lipopeptide. This peptide represents  
 CC an example of a lipopeptide epitope used in the invention and is derived

CC from the human papilloma virus E6 protein. The immunogenic lipopeptide  
 CC micelles are used in vaccines, especially against HIV, hepatitis B virus  
 CC (HBV), papilloma viruses, p53, melanoma or Plasmodium falciparum malaria.  
 XX  
 XX Sequence 23 AA;

Query Match 100.0%; Score 124; DB 20; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLICIVYRDGNPY 22  
 Db 2 RREYDFAFRDLICIVYRDGNPY 23

RESULT 4

AAB95958  
 ID AAB95958 standard; Peptide; 24 AA.

XX AC AAB95958;

XX DT 25-JUN-2001 (first entry)

XX DE HPV 16 E6 protein ffragment SEQ ID 65.

XX KW Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;  
 KW human papillomavirus-associated disease; condyloma; cervical dysplasia;  
 KW cervical dysplasia; major histocompatibility complex; MHC I.

XX OS Human papillomavirus.

XX PN WO200119408-A1.

XX PD 22-MAR-2001.

XX PF 18-SEP-2000; 2000WO-US25559.

XX PR 16-SEP-1999; 99US-0154665.

XX PR 16-SEP-1999; 99US-0398534.

XX PR 09-DEC-1999; 99US-0169846.

XX PR 09-DEC-1999; 99US-0458173.

XX PA (ZYCO-) ZYCO INC.

XX PI Hedley-Mir, Urban-RC, Chicx-RM<sub>2</sub>

XX DR WPI; 2001-265996/27.

XX PT Novel nucleic acids encoding polypeptide polypeptides containing  
 XX multiple epitopes from one or more proteins, useful for treating tumors  
 XX and as vaccines against pathogenic agents

XX PS Claim 42; Page 51; 64pp; English.

XX CC This invention relates to polynucleotides encoding a hybrid polypeptide  
 CC comprising a signal sequence and three segments that are either  
 CC contiguous or separated by a spacer amino acid or spacer peptide. The  
 CC invention specifically details polynucleotides encoding a polypeptide  
 CC peptide where the peptide segments are tumour antigens or a naturally  
 CC occurring protein of a pathogenic agent. The polypeptide peptides exhibit  
 CC antiviral and immunostimulatory activity. The polynucleotide and  
 CC polypeptide peptides are useful for eliciting an immune response in a  
 CC mammal. The polynucleotide and protein are useful as vaccines for  
 CC treating tumours and pathogenic infections. The polynucleotide is also  
 CC useful for preventing or treating human papillomavirus (HPV)-associated  
 CC diseases, particularly exophytic condyloma, flat condyloma, cervical  
 CC cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV  
 CC infection, cervical dysplasia, high grade squamous intraepithelial  
 CC lesions, and anal HPV infection. The polynucleotide and polypeptide are  
 CC useful for generating or enhancing prophylactic or therapeutic immune  
 CC response against pathogens, tumours or autoimmune diseases in a  
 CC population of individuals having diverse MHC allotypes, as positive  
 CC controls in T cell stimulation assays in vitro, and as tools to

CC understand processing of epitopes within cells. Peptides  
 CC AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major  
 CC histocompatibility complex I (MHC I) associated tumour and pathogen  
 CC antigens. The peptides can be used as part of the polypeptide proteins of  
 CC the invention. Also included are examples of the polypeptide proteins  
 CC represented by AAB96050 - AAB96052, and localisation signal peptides  
 CC AAB96038 - AAB96043 and AAB96049 which can be used in the construction of  
 CC the polypeptide peptides.

XX Sequence 24 AA;

Query Match 100.0%; Score 124; DB 22; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLICIVYRDGNPY 22  
 Db 3 RREYDFAFRDLICIVYRDGNPY 24

RESULT 5

AAB31017  
 ID AAB31017 standard; peptide; 32 AA.

XX AC AAB31017;

XX DT 02-APR-2001 (first entry)

XX DE Polyepitopic fragment from HPV E6 protein, comprising residues 46-77.

XX KW Polyepitopic fragment; Nef protein; HIV; human leukocyte antigen; HLA;  
 KW immune response; cytolysis; cytotoxic T cell; CTL; cytokine secretion;  
 KW interleukin-2; IL-2; IL-4; gamma-interferon; vaccine; HPV; E6 protein;  
 KW acquired immune deficiency syndrome.

XX OS Human papillomavirus.

XX PN FR2794370-A1.

XX PD 08-DEC-2000.

XX PF 03-JUN-1999; 99FR-0007012.

XX PR 03-JUN-1999; 99FR-0007012.

XX PA (BIOV-) BIOVECTOR THERAPEUTICS SA.

XX PI (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX DR Choppin J, Bourgault VI, Guillet JG, Connan F, Ferries E;

XX WPI; 2001-064174/08.

XX PT New polypeptopic fragments from Nef protein of the human immune  
 XX deficiency virus, useful for treatment or prevention of acquired immune  
 XX deficiency syndrome

XX PS Disclosure; Page 4; 24pp; French.

XX CC The present sequence represents a polypeptopic fragment from the HPV E6  
 CC protein. The specification describes polypeptopic fragments from the Nef  
 CC protein of human immune deficiency virus (HIV). The Nef peptides bind  
 CC stably to human leukocyte antigen (HLA) type molecules. The Nef peptides  
 CC induce a specific immune response. Particularly, they induce cytolysis,  
 CC by cytotoxic T cells (CTL), of cells that express Nef associated with  
 CC appropriate HLA molecules and induce secretion of cytokines (particularly  
 CC interleukin (IL)-2 and IL-4, and gamma-interferon by these CTL. The Nef  
 CC peptides, their derivatives, nucleic acids encoding them and specific  
 CC antibodies are used, in compositions or vaccines, to treat or prevent  
 CC diseases associated with human immunodeficiency virus (HIV) infection,  
 CC specifically acquired immune deficiency syndrome. The antibodies are  
 CC also useful for in vitro diagnosis of these diseases.

XX SQ Sequence 32 AA;

Query Match 100.0%; Score 124; DB 22; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCLIVYRDGNPY 22  
 DB 1 RREYDFAFRDLCLIVYRDGNPY 22

## RESULT 6

AAB31109

ID AAB31109 standard; peptide; 32 AA.

XX AAB31109;

AC AAB31109;

XX AAB31109;

DT 02-APR-2001 (first entry)

XX AAB31109;

DE A polyepitopic fragment of HPV E6 protein comprising residues 46-77.

XX Polyepitopic peptide; p53; cancer; human leukocyte antigen; HLA; HPV;

KW immune response; cytotoxic T cell; CTL; cytokine secretion;

KW interleukin-2; IL-2; IL-4; gamma-interferon; p53-related cancer;

KW E6 protein.

XX Human papillomavirus.

OS Human papillomavirus.

XX FR2794368-A1.

PN FR2794368-A1.

XX 08-DEC-2000.

PD 07-OCT-1999; 99FR-0012512.

XX 03-JUN-1999; 99FR-0007012.

PR (BIOV-) BIOVECTOR THERAPEUTICS SA.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA Choppin J, Bourgault VI, Gullet JG, Connan F, Ferries E;

XX WPI; 2001-064173/08.

DR New polyepitopic fragments from the p53 protein, useful for treatment

XX or prevention of cancer, e.g. of breast or colon

PT Disclosure; Page 4; 26pp; French.

XX The present sequence represents a polyepitopic fragment of Human

XX papillomavirus (HPV) E6 protein. The specification describes polyepitopic

XX fragments of human p53, which is overexpressed in many types of cancers.

XX The p53 polyepitopic fragments bind stably to human leukocyte antigen

XX (HLA) type molecules. The p53 peptides induce a specific immune response.

XX They induce cytotoxicity, by cytotoxic T cells (CTL), of cells that express

XX the p53 peptides associated with appropriate HLA molecules and induce

XX secretion of cytokines (particularly interleukin (IL)-2 and IL-4, and

XX gamma-interferon) by these CTL. The p53 peptides, derivatives, nucleic

XX acids encoding them and specific antibodies are used, in compositions

XX or vaccines, to treat or prevent diseases p53-related cancers,

XX particularly of breast, colon, lung or bladder.

XX Sequence 32 AA;

SQ Query Match 100.0%; Score 124; DB 22; Length 32;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-13;

XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCLIVYRDGNPY 22

DB 1 RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

ID AAB96050 standard; Peptide; 117 AA.

XX AAB96050;

AC AAB96050;

XX 25-JUN-2001 (first entry)

DT Polyepitope polypeptide SEQ ID 126.

XX Epitope; tumour antigen; antiviral; immunostimulatory; cervical cancer;

XX human papillomavirus-associated disease; condyloma; cervical dysplasia;

XX cervical dysplasia; major histocompatibility complex; MHC I.

XX Synthetic.

XX WO200119408-A1.

XX 22-MAR-2001.

XX 18-SEP-2000; 2000WO-US25559.

XX 16-SEP-1999; 99US-0154665.

XX 16-SEP-1999; 99US-0398534.

XX 09-DEC-1999; 99US-0169846.

XX 09-DEC-1999; 99US-0458173.

XX (ZYCO-) ZYCOS INC.

XX Hedley ML, Urban RC, Chicz RM;

XX WPI; 2001-265996/27.

XX Novel nucleic acids encoding polyepitope polypeptides containing

XX multiple epitopes from one or more proteins, useful for treating tumors

XX and as vaccines against pathogenic agents

XX Disclosure; Fig 2; 64pp; English.

XX This invention relates to polynucleotides encoding a hybrid polypeptide

XX comprising a signal sequence and three segments that are either

XX contiguous or separated by a spacer amino acid or spacer peptide. The

XX invention specifically details polynucleotides encoding a polypeptide

XX peptide where the peptide segments are tumour antigens or a naturally

XX occurring protein of a pathogenic agent. The polynucleotide and

XX polypeptide peptides are useful for eliciting an immune response in a

XX mammal. The polynucleotide and protein are useful as vaccines for

XX treating tumours and pathogenic infections. The polynucleotide is also

XX useful for preventing or treating human papillomavirus (HPV)-associated

XX diseases, particularly exophytic condyloma, flat condyloma, cervical

XX cancer, respiratory papilloma, conjunctival papilloma, genital-tract HPV

XX infection, cervical dysplasia, high grade squamous intraepithelial

XX lesions, and anal HPV infection. The polynucleotide and polypeptide are

XX useful for generating or enhancing prophylactic or therapeutic immune

XX response against pathogens, tumours or autoimmune diseases in a

XX population of individuals having diverse MHC allotypes, as positive

XX controls in T cell stimulation assays in vitro, and as tools to

XX understand processing of epitopes within cells. Peptides

XX AAB95894 - AAB96037 and AAB96044 - AAB96048 represent major

XX histocompatibility complex I (MHC I) associated tumour and pathogen

XX antigens. The peptides can be used as part of the polypeptide proteins of

XX the invention. Also included are examples of the polypeptide proteins

XX represented by AAB96050 - AAB96052, and localisation signal peptides

XX AAB96038 - AAB96043 and AAB96049 which can be used in the construction of

XX the polypeptide peptides.

XX Sequence 117 AA;

SQ Query Match 100.0%; Score 124; DB 22; Length 117;

XX Best Local Similarity 100.0%; Pred. No. 4.8e-13;

XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCLIVYRDGNPY 22

DB 1 RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22

XX RREYDFAFRDLCLIVYRDGNPY 22



Db 23 RREVDFAFRDLCIVRDGNPY 44

# RESULT 8

AA57808  
ID AAY57808 standard; peptide; 151 AA.

AC AAY57808;

XX 20-MAR-2000 (first entry)

XX HPV-16 E6 protein amino acid sequence.

XX Transcriptional adaptor motif; TRAM; TRAM interaction motif; TRIM;  
KW creb binding protein; transcriptional regulation; cytostatic; antiviral;  
KW cell cycle inhibitor; viral transcription inhibitor; cancer; tumour;  
KW viral disease; viral infection; cell cycle; apoptosis; growth arrest.

XX Human papillomavirus.

XX WO9961608-A2.

PN 02-DEC-1999.

XX 26-MAY-1999; 99WO-GB01668.

XX 26-MAY-1998; 98GB-0011303.

PR 05-JAN-1999; 99GB-0000157.

XX (MOLE-) INST MOLECULAR & CELL BIOLOGY.

XX Q'Connor MJ, Zimmermann H;

PI 2000-072620/06.

XX Novel polypeptides cells useful for treating viral disease and cancer

XX Example 2; Fig 9; 73pp; English.

XX The present invention describes a polypeptide comprising a  
CC transcriptional adaptor motif (TRAM) or a TRAM-interaction motif  
CC (TRIM) (I). (I) can be used for identifying compounds (II) comprising a  
CC polypeptide capable of disrupting an interaction between a TRAM sequence  
CC and/or a TRIM sequence. Polypeptides which bind to a TRAM sequence or a  
CC TRIM sequence are identified by incubating, the polypeptide with (I) and  
CC determining if the polypeptide interacts with (I). Compound (II) is  
CC useful for preparing a pharmaceutical composition and for disrupting an  
CC interaction between TRAM sequence and TRIM sequence in vitro, thereby  
CC inhibiting viral transcription or cell cycle progression in mammalian  
CC cells especially cancer cell. Compounds which disrupt interaction  
CC between TRIM/TRAM containing polypeptides can be used therapeutically to  
CC prevent or treat viral diseases and tumours. The polypeptides reduce  
CC susceptibility of cells to viral infection and regulate cell cycle  
CC including apoptosis and growth arrest and can be used to produce  
CC antibodies against the TRIM or TRAM sequences. HPV types associated  
CC with high risk or low risk of cervical cancer can be distinguished  
CC based on the ability of E6 polypeptides to bind to creb binding protein  
CC (CBP) TRAM sequence. The present sequence represents a HPV-16 E6 protein  
CC amino acid sequence from an example from the present invention.

XX Sequence 151 AA;

Query Match 100.0%; Score 124; DB 21; Length 151;

Best Local Similarity 100.0%; Pred. No. 6.4e-13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREVDFAFRDLCIVRDGNPY 22

DB 39 RREVDFAFRDLCIVRDGNPY 60

# RESULT 9

AAR22766

ID AAR22766 standard; peptide; 158 AA.

XX AAR22766;

XX 21-SEP-1992 (first entry)

XX HPV E6 peptide.

XX Human; papillomavirus; immunogenic; cervical; warts; carcinoma;  
KW cancer.

XX Synthetic.

OS Homo sapiens.

XX WO9205248-A\*

PD 02-APR-1992.

XX 26-SEP-1991; 91WO-US07081.

XX 26-SEP-1990; 90US-0588384.

XX (BRIM ) BRISTOL-MYERS SQUIB.

XX Blake J, Chen L, Hellstrom I, Hellstrom K, Hu S L;  
PI Thomas E K;

XX WPI; 1992-132119/16.

XX Immunogenic peptide(s) derived from E6 or E7 region of HPV16 -  
PT and recombinant cells encoding them, useful in treatment and  
PT prophylaxis of cervical warts or cancer resulting from HPV  
PT infection

XX Disclosure; Fig 7; 81pp; English.

XX The peptide is the sequence of the human papillomavirus HPV 16 E6  
CC nucleoprotein. Peptides corresponding to regions (pref. epitopic  
CC regions) of HPV 16 E6 were synthesised by standard Merrifield  
CC synthesis. Examples of such peptides are E6 1-20, 8-20, 119-134 or  
CC 148-158. Compositions contg. these peptides, antibodies against the  
CC peptides, or recombinant cells contg. the gene encoding the immuno-  
CC genic peptides may be utilised in methods for inhibiting and treating  
CC HPV infection and tumour initiation and progression e.g. in the  
CC prevention or retardation of cervical warts and cervical carcinoma  
CC resulting from HPV infection.  
CC See also AAR22767.

XX Sequence 158 AA;

Query Match 100.0%; Score 124; DB 13; Length 158;

Best Local Similarity 100.0%; Pred. No. 6.8e-13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREVDFAFRDLCIVRDGNPY 22

DB 46 RREVDFAFRDLCIVRDGNPY 67

# RESULT 10

AA582462

ID AAY82462 standard; Protein; 158 AA.

XX AAY82462;

XX 30-JUN-2000 (first entry)

XX Human papillomavirus E6 protein containing two zinc finger motifs.

XX Chelated zinc finger; therapeutic; treatment; prophylaxis; MPV;  
KW mammalian papillomavirus; antiviral; cytostatic; cervical cancer;  
KW lesion; wart.

XX

OS Human papillomavirus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 37..73 /note= "forms a zinc finger motif"  
 FT FT  
 FT Misc-difference 110..146 /note= "forms a zinc finger motif"  
 FT FT  
 XX WO200014063-A1.  
 PN XX  
 PD 16-MAR-2000.  
 XX  
 XX 03-SEP-1999; 99WO-AU00724.  
 PF  
 XX 04-SEP-1998; 98AU-0005733.  
 PR  
 XX 15-JUL-1999; 99AU-0001645.  
 XX  
 PA (MOLE-) INST MOLECULAR & CELL BIOLOGY.  
 PA (HUGH/) HUGHES E J L.  
 XX  
 XX Bernard H, Tan YJ, Beerheide W, Ting AE, Sim MM;  
 PI WPI; 2000-256917/22.  
 DR  
 XX  
 XX Polysulfide and dithionodisulfide agents, useful for the treatment or  
 PT prophylaxis of diseases caused by mammalian papillomavirus, are  
 PT disruptors of a chelated metal cation domain in an MPV gene encoded  
 PT protein.  
 XX  
 XX Disclosure; Fig 1; 78pp; English.  
 PS  
 XX The present invention describes an agent used in the treatment or  
 CC prophylaxis of a disease caused or exacerbated by MPV (mammalian  
 CC papillomavirus) comprising a compound capable of reducing, inhibiting  
 CC or otherwise decreasing the activity of a protein encoded by an MPV  
 CC gene by facilitating disruption of a chelated metal cation domain  
 CC present in the protein. An agent of the present invention can be used  
 CC to treat cervical cancer or its HPV associated precursor lesions or  
 CC other HPV associated cancers and/or warts. The present sequence  
 CC represents a human papillomavirus E6 protein containing two zinc  
 CC finger motifs, as given in the exemplification of the present  
 CC invention.  
 XX  
 XX Sequence 158 AA;  
 SQ  
 Query Match 100.0%; Score 124; DB 21; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-13;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RREYDFAFRDLCIVYRDGNPY 22  
 DB 46 RREYDFAFRDLCIVYRDGNPY 67  
 RESULT 11  
 AAB98420  
 ID AAB98420 standard; Protein; 158 AA.  
 XX  
 XX AAB98420;  
 AC  
 XX 22-AUG-2001 (first entry)  
 DT  
 DE Human papillomavirus protein HPV16 E6.  
 XX  
 XX Human papillomavirus; human leukocyte antigen; HLA; immune response;  
 KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
 KW neoplastic growth; antiviral.  
 XX  
 XX Human papillomavirus.  
 OS  
 XX WO200141799-A1.  
 PN  
 XX 14-JUN-2001.  
 PD

XX 11-DEC-2000; 2000WO-US33549.  
 XX  
 XX 10-DEC-1999; 99US-0172705.  
 PR  
 XX 15-AUG-2000; 2000US-0641528.  
 XX  
 XX (EPIM-) EPIMUNE INC.  
 XX  
 XX Sette A, Sidney J, Southwood S, Chesnut R, Cellis E, Grey HM;  
 PI WPI; 2001-381497/40.  
 XX  
 XX An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
 PT treating HPV infections -  
 PT  
 XX Disclosure; Page 20-21; 756pp; English.  
 PS  
 XX The present invention describes an isolated prepared human papillomavirus  
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses,  
 CC specifically in the treatment or prophylaxis of HPV infection, in persons  
 CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
 CC The peptides can also be used in a tetramer staining assay to assess  
 CC peripheral blood mononuclear cells for the presence of antigen-specific  
 CC CTLs following exposure to a pathogen or immunogen, and as reagents to  
 CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
 CC The vaccine compositions are useful for removing warts or treating HPV  
 CC infections. The epitopes for inclusion in an epitope-base vaccine may  
 CC be selected from conserved regions of viral or tumour-associated  
 CC antigens, which reduces the likelihood of escape mutants, also  
 CC immunosuppressive epitopes that may be present in whole antigens can be  
 CC avoided with the use of epitope-base vaccines. An additional advantage  
 CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
 CC the composition of the epitopes achieving enhanced immunogenicity, the  
 CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
 CC to AAB98477 represent polypeptide sequences used in the exemplification  
 CC of the present invention.  
 XX  
 XX Sequence 158 AA;  
 SQ  
 Query Match 100.0%; Score 124; DB 22; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-13;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RREYDFAFRDLCIVYRDGNPY 22  
 DB 46 RREYDFAFRDLCIVYRDGNPY 67  
 RESULT 12  
 AAB35741  
 ID AAB35741 standard; Protein; 162 AA.  
 XX  
 XX AAB35741;  
 AC  
 XX 16-FEB-1998 (first entry)  
 DT  
 XX Human papillomavirus type 16 E6 protein.  
 DE  
 XX Complete genome; circular; human papillomavirus type 16; HPV16 E6;  
 KW cervical dysplasia; cervical cancer; cervical smear.  
 KW  
 XX Human papillomavirus type 16.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 90 /note= "Mutated to Val in the variant"  
 FT  
 FT Misc-difference 159 /note= "End of protein sequence even though 3 amino  
 FT acid residues are given following on"  
 FT  
 XX

```

PN US5679509-A.
XX
PD 21-OCT-1997.
XX
PF 28-SEP-1993; 93US-0127906.
XX
PR 30-SEP-1994; 94US-0316239.
PR 28-SEP-1993; 93US-0127906.
XX
PA (UYNE-) UNIV NEW MEXICO STATE.
XX
PI Parmenter CA, Wheeler CM;
XX
DR WPI; 1997-525714/48.
DR N-PSDB; AAT94723.
XX
PT Evaluating risk of cervical dysplasia or cervical cancer - by
PT detecting variant form of human papilloma virus 16
XX
PS Claim 7; Column 23-24; 33pp; English.
XX
CC Methods have been developed for distinguishing a subset of human
CC papilloma virus (HPV) that is associated with an increased risk of
CC developing cervical dysplasia or cervical cancer. The methods involve:
CC (1) preparing a cervical sample to expose any HPV-16 E6 gene in the
CC sample and determining if the base at position 350 of the E6 gene
CC (see AAT94723 and AAT94724 for comparison) is T or G, where the presence
CC of G at position 350 is associated with an increased risk of developing
CC cervical dysplasia or cervical cancer; and (2) preparing a cervical
CC sample to expose any HPV-16 E6 protein in the sample and determining
CC if the amino acid at position 83 of the protein (see position 90 in
CC AAW35741 and AAW35742 for comparison) is Val or Leu, where the presence
CC of Val at position 83 that is associated with an increased risk of
CC developing cervical dysplasia or cervical cancer. The present sequence
CC represents the reference protein sequence for HPV-16 E6. The 350G
CC variant correlates well with Pap scores: 350T:350G ratios among 45
CC HPV16 samples were 10:4 for negative Pap scores; 4:2 for CIN I, 1:6 for
CC CIN II; 2:9 for CIN III; 0:3 for cancer.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 124; DB 18; Length 162;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
DB |||||
46 RREYDFAFRDLCIVYRDGNPY 67

RESULT 13
AAW35742
ID AAW35742 standard; Protein; 162 AA.
XX
AC AAW35742;
XX
DT 16-FEB-1998 (first entry)
XX
DE Human papillomavirus type 16 E6 protein variant.
XX
KW Complete genome; circular; human papillomavirus type 16; HPV16 E6;
KW cervical dysplasia; cervical cancer; cervical smear.
XX
OS Human papillomavirus type 16.
XX
FH Key Location/Qualifiers
FT Misc-difference 90
FT /note= "Mutated from Leu in the reference sequence
FT (AAW35741)"
FT
FT Misc-difference 159
FT /note= "End of protein sequence even though 3 amino
FT acid residues are given following on"
XX

PN US5679509-A.
XX
PD 21-OCT-1997.
XX
PF 28-SEP-1993; 93US-0127906.
XX
PR 30-SEP-1994; 94US-0316239.
PR 28-SEP-1993; 93US-0127906.
XX
PA (UYNE-) UNIV NEW MEXICO STATE.
XX
PI Parmenter CA, Wheeler CM;
XX
DR WPI; 1997-525714/48.
DR N-PSDB; AAT94742.
XX
PT Evaluating risk of cervical dysplasia or cervical cancer - by
PT detecting variant form of human papilloma virus 16
XX
PS Claim 7; Column 23-26; 33pp; English.
XX
CC Methods have been developed for distinguishing a subset of human
CC papilloma virus (HPV) that is associated with an increased risk of
CC developing cervical dysplasia or cervical cancer. The methods involve:
CC (1) preparing a cervical sample to expose any HPV-16 E6 gene in the
CC sample and determining if the base at position 350 of the E6 gene
CC (see AAT94723 and AAT94724 for comparison) is T or G, where the presence
CC of G at position 350 is associated with an increased risk of developing
CC cervical dysplasia or cervical cancer; and (2) preparing a cervical
CC sample to expose any HPV-16 E6 protein in the sample and determining
CC if the amino acid at position 83 of the protein (see position 90 in
CC AAW35741 and AAW35742 for comparison) is Val or Leu, where the presence
CC of Val at position 83 that is associated with an increased risk of
CC developing cervical dysplasia or cervical cancer. The present sequence
CC represents the variant protein sequence for HPV-16 E6. The 350G
CC variant correlates well with Pap scores: 350T:350G ratios among 45
CC HPV16 samples were 10:4 for negative Pap scores; 4:2 for CIN I, 1:6 for
CC CIN II; 2:9 for CIN III; 0:3 for cancer.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 124; DB 18; Length 162;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
DB |||||
46 RREYDFAFRDLCIVYRDGNPY 67

RESULT 14
AAR97563
ID AAR97563 standard; Protein; 172 AA.
XX
AC AAR97563;
XX
DT 11-JAN-1997 (first entry)
XX
DE Human papilloma virus E6/E7 protein variant.
XX
KW Human papilloma virus; E6; E7; deletion mutant; HPV;
KW immune response; humoral immune response; cellular immune response;
KW vaccine.
XX
OS Human papilloma virus.
XX
PN WO9619496-A1.
XX
PD 27-JUN-1996.
XX
PF 20-DEC-1995; 95WO-AU00868.
XX
PR 20-DEC-1994; 94AU-0000157.
XX

```

```
XX (CSLC-) CSL LTD.
PA (UYQU ) UNIV QUEENSLAND.
XX
XX Cox J, Edwards SJ, Frazer I, Webb EA;
XX
XX WPI; 1996-309518/31.
DR N-PSDB; AAT31835.
XX
XX Vaccine variants of human papilloma virus antigens - contain
PT variants of E6 and/or E7 protein, pref. deletion mutants, and are
PT used to treat or prevent HPV infection
XX
XX Example 3; Page 18; 37pp; English.
XX
XX A variant of the human papilloma virus (HPV) E6 or E7 protein which
CC elicits a humoral and/or cellular immune response against HPV can be
CC used in vaccines against HPV or to treat HPV infection. The variant
CC is preferably a deletion mutant comprising at least half, and
CC preferably two-thirds of full length E6 or E7 protein starting from
CC the N- or C-terminal, or is a full length E6 moiety fused to a full
CC length E7 moiety. The variant optionally has a linkage moiety and a
CC foreign protein or peptide which facilitates the purification of,
CC and enhances the immunogenicity of, the fusion protein. This
CC sequence is a fusion protein of the C-terminal end of E7 and the
CC N-terminal end of E6.
XX
XX Sequence 172 AA;
SQ
Query Match 100.0%; Score 124; DB 17; Length 172;
Best Local Similarity 100.0%; Pred. No. 7.5e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||
Db 115 RREYDFAFRDLCIVYRDGNPY 136

RESULT 15
AAR63865
ID AAR63865 standard; Protein; 188 AA.
XX
XX AAR63865;
AC
XX
XX 28-JUN-1995 (first entry)
DT
XX
XX HPV16 E6/E7 proteins.
DE
XX
XX HPV; HPV16; E6 protein; E7 protein; diagnosis; cervical dysplasia;
KW cervix cancer.
XX
XX Human papillomavirus strain 16.
OS
XX
XX Key Location/Qualifiers
FH 1..158
FT Protein /label= E6_protein
FT Protein 159..188
FT /label= E7_protein
XX
XX W09426934-A.
PN
XX
XX 24-NOV-1994.
PD
XX
XX 06-MAY-1994; 94WO-US05085.
PF
XX
XX 06-MAY-1993; 93US-0058920.
PR
XX
XX (BAXT ) BAXTER DIAGNOSTICS INC.
PA
XX
XX Brown JT;
PI
XX
XX WPI; 1995-006821/01.
DR
XX P-PSDB; AAR63865.
```

```
XX Human papilloma virus detection assay - by amplification using
PT self sustained sequence replication and hybridisation with a
PT detector probe
XX
XX Disclosure; Page 24-26; 79pp; English.
PS
XX
XX The sequences of the E6 and E7 polypeptide-encoding regions of human
CC papillomavirus (HPV) 16 and 18 are given in AAR63865-71 and the
CC encoded proteins in AAR63865-66, respectively. Probes and primers
CC based on these sequences were used for HPV infection diagnosis;
CC expression of E6 and E7 is diagnostic for cervical cancer or pre-
CC malignant states.
XX
XX Sequence 188 AA;
SQ
Query Match 100.0%; Score 124; DB 16; Length 188;
Best Local Similarity 100.0%; Pred. No. 8.3e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||
Db 46 RREYDFAFRDLCIVYRDGNPY 67

Search completed: July 3, 2003, 16:41:15
Job time : 71 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 16:41:21 ; Search time 308 Seconds  
(without alignments)  
46.052 Million cell updates/sec

Title: US-09-980-523A-6

Perfect score: 124

Sequence: 1 REVYDFAFRLCIYVDGNPY 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	124	100.0	22	23	US-09-980-523A-6
2	124	100.0	23	19	US-09-555-780-125
3	124	100.0	23	19	US-09-555-780B-125
4	124	100.0	23	20	US-09-601-729-276
5	124	100.0	24	1	PCT-US00-25559-65
6	124	100.0	24	20	US-09-664-225-65

7	124	100.0	117	1	PCT-US00-25559-126	Sequence 126, App
8	124	100.0	117	20	US-09-664-225-126	Sequence 126, App
9	124	100.0	151	15	US-09-177-390-6	Sequence 6, Appl
10	124	100.0	151	21	US-09-701-080A-18	Sequence 20, Appl
11	124	100.0	151	27	US-60-306-809-20	Sequence 13, Appl
12	124	100.0	151	27	US-60-415-929-13	Sequence 31, Appl
13	124	100.0	158	1	PCT-US00-33549-31	Sequence 2, Appl
14	124	100.0	158	23	US-09-980-523A-2	Sequence 157, App
15	124	100.0	236	1	PCT-US00-25559-157	Sequence 157, App
16	124	100.0	236	20	US-09-664-225-157	Sequence 158, App
17	124	100.0	237	1	PCT-US00-25559-158	Sequence 158, App
18	124	100.0	237	20	US-09-664-225-158	Sequence 1, Appl
19	124	100.0	243	18	US-09-462-993-1	Sequence 1, Appl
20	124	100.0	248	27	US-60-415-929-1	Sequence 3, Appl
21	124	100.0	248	27	US-60-415-929-3	Sequence 5, Appl
22	124	100.0	248	27	US-60-415-929-5	Sequence 7, Appl
23	124	100.0	248	27	US-60-415-929-7	Sequence 9, Appl
24	124	100.0	248	27	US-60-415-929-9	Sequence 11, Appl
25	124	100.0	248	27	US-60-415-929-11	Sequence 160, App
26	124	100.0	261	1	PCT-US00-25559-160	Sequence 160, App
27	124	100.0	261	20	US-09-664-225-160	Sequence 4, Appl
28	124	100.0	273	19	US-09-581-976-4	Sequence 4, Appl
29	124	100.0	273	24	US-10-000-903-4	Sequence 10, Appl
30	124	100.0	292	19	US-09-581-976-10	Sequence 6, Appl
31	124	100.0	292	24	US-10-000-903-10	Sequence 6, Appl
32	124	100.0	371	19	US-09-581-976-6	Sequence 14, Appl
33	124	100.0	371	24	US-10-000-903-6	Sequence 14, Appl
34	124	100.0	390	19	US-09-581-976-14	Sequence 161, App
35	124	100.0	390	24	US-10-000-903-14	Sequence 58, Appl
36	97	78.2	20	3	US-07-949-836C-161	Sequence 45, Appl
37	93	75.0	149	1	PCT-US00-33549-58	Sequence 51, Appl
38	72	58.1	149	1	PCT-US00-33549-45	Sequence 38, Appl
39	72	58.1	158	1	PCT-US00-33549-51	Sequence 4, Appl
40	71	57.3	158	5	US-08-176-937A-4	Sequence 21, Appl
41	71	57.3	158	19	US-09-581-976-21	Sequence 21, Appl
42	71	57.3	278	24	US-10-000-903-21	Sequence 23, Appl
43	71	57.3	278	24	US-10-000-903-21	Sequence 23, Appl
44	71	57.3	383	19	US-09-581-976-23	Sequence 23, Appl
45	71	57.3	383	24	US-10-000-903-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-09-980-523A-6  
; Sequence 6, Application US/09980523A  
; GENERAL INFORMATION:  
; APPLICANT: CHOPPIN, JEANNINE  
; APPLICANT: BOURGALT VILLADA, ISABELLE  
; APPLICANT: GUILLET, JEAN-GERARD  
; APPLICANT: CONNAN, FRANCINE  
; APPLICANT: FERRIES, ESTELLE  
; TITLE OF INVENTION: POLYPEPTIC PROTEIN FRAGMENTS OF THE E6 AND E7  
; TITLE OF INVENTION: PROTEINS OF HPV, THEIR PRODUCTION AND THEIR USE  
; FILE REFERENCE: WOBI AO INS  
; CURRENT APPLICATION NUMBER: US/09/980, 523A  
; CURRENT FILING DATE: 2002-04-29  
; PRIOR APPLICATION NUMBER: PCT/FR00/01513  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: FR 99/07012  
; PRIOR FILING DATE: 1999-06-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Human Papillomavirus  
US-09-980-523A-6

Query Match 100.0%; Score 124; DB 23; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.4e-12;

```
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||
Db 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||

RESULT 2
US-09-555-780-125
; Sequence 125, Application US/09555780
; GENERAL INFORMATION:
; APPLICANT: INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE
; APPLICANT: M
; APPLICANT: (C.N.)
; APPLICANT: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
; APPLICANT: INSTITUT PASTEUR DE LILLE
; APPLICANT: GRAS-MASSE, Helene
; APPLICANT: BOSSUS, Marc
; APPLICANT: LIPPENS, Guy
; APPLICANT: WIERUSZESKI, Jean-Michel
; APPLICANT: TARTAR, Andre
; APPLICANT: GUILLET, Jean-Gerard
; APPLICANT: BOURGAULT-VILLADA, Isabelle
; TITLE OF INVENTION: Mixed lipopeptide micelles for inducing an immune
; TITLE OF INVENTION: response and their therapeutic uses.
; FILE REFERENCE: Lipopeptides-INSERM
; CURRENT APPLICATION NUMBER: US/09/555,780
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: FR9715246
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 125
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Human papillomavirus
US-09-555-780-125

Query Match 100.0%; Score 124; DB 19; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||
Db 2 RREYDFAFRDLCIVYRDGNPY 23
    |||||

RESULT 3
US-09-555-780B-125
; Sequence 125, Application US/09555780B
; GENERAL INFORMATION:
; APPLICANT: INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE M
; APPLICANT: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (C.N.)
; APPLICANT: INSTITUT PASTEUR DE LILLE
; APPLICANT: GRAS-MASSE, H,ISne
; APPLICANT: BOSSUS, Marc
; APPLICANT: LIPPENS, Guy
; APPLICANT: WIERUSZESKI, Jean-Michel
; APPLICANT: TARTAR, Andr,
; APPLICANT: GUILLET, Jean-G.rard
; APPLICANT: BOURGAULT-VILLADA, Isabelle
; TITLE OF INVENTION: Mixed lipopeptide micelles for inducing an immune
; TITLE OF INVENTION: response and their therapeutic uses.
; FILE REFERENCE: Lipopeptides-INSERM
; CURRENT APPLICATION NUMBER: US/09/555,780B
; CURRENT FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: FR9715246
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 285
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 125
; LENGTH: 23
; TYPE: PRT
```

```
; ORGANISM: Human papillomavirus
US-09-555-780B-125

Query Match 100.0%; Score 124; DB 19; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||
Db 2 RREYDFAFRDLCIVYRDGNPY 23
    |||||

RESULT 4
US-09-601-729-276
; Sequence 276, Application US/09601729
; GENERAL INFORMATION:
; APPLICANT: THIAM, KADER
; APPLICANT: AURIAULT, CLAUDE
; APPLICANT: GRAS-MASSE, HELENE
; APPLICANT: LOING, ESTELLE
; APPLICANT: VERWAERDE, CLAUDIE
; APPLICANT: GUILLET, JEAN GERARD
; TITLE OF INVENTION: LIPOPEPTIDES CONTAINING AN INTERFERON FRAGMENT AND USES
; TITLE OF INVENTION: THEREOF IN PHARMACEUTICAL COMPOSITIONS
; FILE REFERENCE: USB-97-AU-IN
; CURRENT APPLICATION NUMBER: US/09/601,729
; CURRENT FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: PCT/FR99/00259
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 98 01439
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 281
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 276
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-601-729-276

Query Match 100.0%; Score 124; DB 20; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22
    |||||
Db 2 RREYDFAFRDLCIVYRDGNPY 23
    |||||

RESULT 5
PCT-US00-25559-65
; Sequence 65, Application PC/TUS0025559
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDE POLYPEPTIDES
; FILE REFERENCE: 08191-013W01
; CURRENT APPLICATION NUMBER: PCT/US00/25559
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/25559
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/169,846
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: US 60/154,665
; PRIOR FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 65
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Human Papilloma virus
PCT-US00-25559-65
```

Query Match 100.0%; Score 124; DB 1; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 3 RREYDFAFRDLCIVYRDGNPY 24

RESULT 6  
US-09-664-225-65  
; Sequence 65, Application US/09664225  
; GENERAL INFORMATION:  
; APPLICANT: Hedley, Mary Lynne  
; APPLICANT: Urban, Robert G.  
; APPLICANT: Chiciz, Roman M.  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDE POLYPEPTIDES  
; FILE REFERENCE: 08191-013001  
; CURRENT APPLICATION NUMBER: US/09/664, 225  
; CURRENT FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/169,846  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: US 60/154,665  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 65  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Human Papilloma virus  
US-09-664-225-65

Query Match 100.0%; Score 124; DB 20; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.8e-12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 3 RREYDFAFRDLCIVYRDGNPY 24

RESULT 7  
PCT-US00-25559-126  
; Sequence 126, Application PC/TUS0025559  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDE POLYPEPTIDES  
; FILE REFERENCE: 08191-013001  
; CURRENT APPLICATION NUMBER: PCT/US00/25559  
; CURRENT FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US00/25559  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US 60/169,846  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: US 60/154,665  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 126  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial fusion sequence  
PCT-US00-25559-126

Query Match 100.0%; Score 124; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||

Db 23 RREYDFAFRDLCIVYRDGNPY 44

RESULT 8  
US-09-664-225-126  
; Sequence 126, Application US/09664225  
; GENERAL INFORMATION:  
; APPLICANT: Hedley, Mary Lynne  
; APPLICANT: Urban, Robert G.  
; APPLICANT: Chiciz, Roman M.  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDE POLYPEPTIDES  
; FILE REFERENCE: 08191-013001  
; CURRENT APPLICATION NUMBER: US/09/664, 225  
; CURRENT FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/169,846  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: US 60/154,665  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 126  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial fusion sequence  
US-09-664-225-126

Query Match 100.0%; Score 124; DB 20; Length 117;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 23 RREYDFAFRDLCIVYRDGNPY 44

RESULT 9  
US-09-177-390-6  
; Sequence 6, Application US/09177390  
; GENERAL INFORMATION:  
; APPLICANT: Schuler, Gerold  
; APPLICANT: N.V. Antwerps Innovatiecentrum  
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear  
; TITLE OF INVENTION: Polynucleotides by Electroporation  
; FILE REFERENCE: 021505wo/JH/ml  
; CURRENT APPLICATION NUMBER: US/09/177,390  
; CURRENT FILING DATE: 2002-06-20  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Human papillomavirus type 16  
US-09-177-390-6

Query Match 100.0%; Score 124; DB 15; Length 151;  
Best Local Similarity 100.0%; Pred. No. 3.4e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 39 RREYDFAFRDLCIVYRDGNPY 60

RESULT 10  
US-09-701-080A-18  
; Sequence 18, Application US/09701080A  
; GENERAL INFORMATION:  
; APPLICANT: O'CONNOR, MARK J.  
; APPLICANT: ZIMMERMAN, HOLGER  
; TITLE OF INVENTION: POLYPEPTIDES FROM CREB BINDING PROTEIN AND RELATED PROTEIN P30  
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION

; FILE REFERENCE: 117-328  
; CURRENT APPLICATION NUMBER: US/09/701.080A  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: GB 9811303.8  
; PRIOR FILING DATE: 1998-05-26  
; PRIOR APPLICATION NUMBER: GB 9900157.0  
; PRIOR FILING DATE: 1999-01-05  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Human papillomavirus  
US-09-701-080A-18

Query Match 100.0%; Score 124; DB 21; Length 151;  
Best Local Similarity 100.0%; Pred. No. 3.4e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 39 RREYDFAFRDLCIVYRDGNPY 60

## RESULT 11

US-60-306-809-20  
; Sequence 20, Application US/60306809  
; GENERAL INFORMATION:  
; APPLICANT: SASTRY, K. JAGANNADHA  
; APPLICANT: TORTOLERO-LUNA, GUILLERMO  
; APPLICANT: FOLLEN, MICHELE  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO HPV-ASSOCIATED  
; TITLE OF INVENTION: PRE-CANCEROUS AND CANCEROUS GROWTHS, INCLUDING CIN  
; FILE REFERENCE: UTSC:560USP1  
; CURRENT APPLICATION NUMBER: US/60/306.809  
; CURRENT FILING DATE: 2001-07-20  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Human papillomavirus  
US-60-306-809-20

Query Match 100.0%; Score 124; DB 27; Length 151;  
Best Local Similarity 100.0%; Pred. No. 3.4e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 39 RREYDFAFRDLCIVYRDGNPY 60

## RESULT 12

US-60-415-929-13  
; Sequence 13, Application US/60415929  
; GENERAL INFORMATION:  
; APPLICANT: Cassetti, Maria  
; TITLE OF INVENTION: HUMAN PAPILLOMAVIRUS POLYPEPTIDES AND IMMUNOGENIC COMPOSITIONS  
; FILE REFERENCE: 0630/OM137-US0  
; CURRENT APPLICATION NUMBER: US/60/415.929  
; CURRENT FILING DATE: 2002-10-03  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Human papillomavirus type 16  
US-60-415-929-13

Query Match 100.0%; Score 124; DB 27; Length 151;  
Best Local Similarity 100.0%; Pred. No. 3.4e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 39 RREYDFAFRDLCIVYRDGNPY 60

## RESULT 13

PCT-US00-33549-31  
; Sequence 31, Application PC/TUS0033549  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Human  
; TITLE OF INVENTION: Papillomavirus Using Peptide and Nucleic Acid  
; TITLE OF INVENTION: Compositions  
; FILE REFERENCE: 018623-016110PC  
; CURRENT APPLICATION NUMBER: PCT/US00/33549  
; CURRENT FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR APPLICATION NUMBER: US 09/641,528  
; PRIOR FILING DATE: 2000-08-15  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 31  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Human papillomavirus type 16  
PCT-US00-33549-31

Query Match 100.0%; Score 124; DB 1; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.6e-11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCIVYRDGNPY 22  
|||||  
Db 46 RREYDFAFRDLCIVYRDGNPY 67

## RESULT 14

US-09-980-523A-2  
; Sequence 2, Application US/09980523A  
; GENERAL INFORMATION:  
; APPLICANT: CHOPPIN, JEANNINE  
; APPLICANT: BOURGAULT VILLADA, ISABELLE  
; APPLICANT: GUILLET, JEAN-GERARD  
; APPLICANT: CONNAN, FRANCINE  
; APPLICANT: FERRIES, ESTELLE  
; TITLE OF INVENTION: POLYPEPTIC PROTEIN FRAGMENTS OF THE E6 AND E7  
; TITLE OF INVENTION: PROTEINS OF HPV, THEIR PRODUCTION AND THEIR USE  
; TITLE OF INVENTION: PARTICULARLY IN VACCINATION  
; FILE REFERENCE: WO01 AO INS  
; CURRENT APPLICATION NUMBER: US/09/980.523A  
; CURRENT FILING DATE: 2002-04-29  
; PRIOR APPLICATION NUMBER: PCT/FR00/01513  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: FR 99/07012  
; PRIOR FILING DATE: 1999-06-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Human Papillomavirus  
US-09-980-523A-2

Query Match 100.0%; Score 124; DB 23; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.6e-11;



Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGPNY 22  
 Db 46 RREYDFAFRDLCIVYRDGPNY 67

RESULT 15

PCT-US00-25559-157  
 ; Sequence 157, Application PC/TUS0025559  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZYCOS INC.  
 ; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDE POLYPEPTIDES  
 ; FILE REFERENCE: 08191-013WO1  
 ; CURRENT APPLICATION NUMBER: PCT/US00/25559  
 ; CURRENT FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: PCT/US00/25559  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US 60/169,846  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: US 60/154,665  
 ; PRIOR FILING DATE: 1999-09-16  
 ; NUMBER OF SEQ ID NOS: 163  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 157  
 ; LENGTH: 236  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Artificial fusion sequence  
 PCT-US00-25559-157

Query Match 100.0%; Score 124; DB 1; Length 236;

Best Local Similarity 100.0%; Pred No. 5.5e-11;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGPNY 22  
 Db 23 RREYDFAFRDLCIVYRDGPNY 44

Search completed: July 3, 2003, 16:49:44  
 Job time : 309 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 6, 2003, 13:05:29 ; Search time 1127 Seconds  
(without alignments)  
568.112 Million cell updates/sec

Title: US-09-980-523A-6  
Perfect score: 124  
Sequence: 1 RREYDFAFRDLICIVRDGPNY 22

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/USG9980523/runat\_03072003\_164406\_4233/app-query.fasta\_1.199  
-DB=GenEmbl -QMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=USG9980523 @CGN\_1.1.3745 @runat\_03072003\_164406\_4233 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGUEURY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	124	100.0	66	6	AX057042 Sequence
2	124	100.0	298	14	AF404704 Human pap
3	124	100.0	421	14	S51110 orf E6 (hum
4	124	100.0	451	14	AF404696 Human pap
5	124	100.0	451	14	AF404699 Human pap
6	124	100.0	451	14	AF404700 Human pap
7	124	100.0	451	14	AF404701 Human pap
8	124	100.0	451	14	AF404703 Human pap
9	124	100.0	456	6	ARI167393 Sequence
10	124	100.0	456	6	ARI177943 Sequence
11	124	100.0	456	14	AF327851 Human pap
12	124	100.0	456	14	HPU34107 Human papil
13	124	100.0	456	14	HPU34108 Human papil
14	124	100.0	456	14	HPU34109 Human papil
15	124	100.0	456	14	HPU34110 Human papil
16	124	100.0	456	14	HPU34111 Human papil
17	124	100.0	456	14	HPU34112 Human papil
18	124	100.0	456	14	HPU34113 Human papil
19	124	100.0	456	14	HPU34114 Human papil
20	124	100.0	456	14	HPU34115 Human papil
21	124	100.0	456	14	HPU34116 Human papil
22	124	100.0	456	14	HPU34117 Human papil
23	124	100.0	456	14	HPU34118 Human papil
24	124	100.0	456	14	HPU34119 Human papil
25	124	100.0	456	14	HPU34120 Human papil
26	124	100.0	456	14	HPU34121 Human papil
27	124	100.0	456	14	HPU34122 Human papil
28	124	100.0	456	14	HPU34123 Human papil
29	124	100.0	456	14	HPU34124 Human papil
30	124	100.0	456	14	HPU34125 Human papil
31	124	100.0	456	14	HPU34126 Human papil
32	124	100.0	456	14	HPU34127 Human papil
33	124	100.0	456	14	HPU34128 Human papil
34	124	100.0	456	14	HPU34129 Human papil
35	124	100.0	456	14	HPU34130 Human papil
36	124	100.0	456	14	HPU34131 Human papil
37	124	100.0	456	14	HPU34132 Human papil
38	124	100.0	456	14	HPU34133 Human papil
39	124	100.0	456	14	HPU34134 Human papil
40	124	100.0	456	14	HPU34135 Human papil
41	124	100.0	458	14	AF003013 Human pap
42	124	100.0	458	14	AF003014 Human pap
43	124	100.0	458	14	AF003015 Human pap
44	124	100.0	458	14	AF003016 Human pap
45	124	100.0	458	14	AF003017 Human pap

ALIGNMENTS

RESULT 1

AX057042  
LOCUS AX057042 66 bp DNA linear PAT 17-JAN-2001  
DEFINITION Sequence 5 from Patent WO075336.  
ACCESSION AX057042  
VERSION AX057042.1 GI:12309883  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial construct.  
REFERENCE  
AUTHORS Choppin,J., Bourgault Villada,I., Guillet,J.G., Connan,F. and  
Ferries,E.  
TITLE 1 (bases 1 to 66)  
Choppin,J., Bourgault Villada,I., Guillet,J.G., Connan,F. and  
Ferries,E.  
JOURNAL Patent: WO 0075336-A 5 14-DEC-2000;  
Biovector Therapeutics S.A. (FR); Institut National de la Sante et  
de la Recherche Medicale (INSERM) (FR)  
FEATURES  
source Location/Qualifiers  
1..66  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="fragment de la sequence codant pour E6 de HPV et  
sequence peptidique correspondante"  
CDS  
1..>66  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC22412.1"  
/db\_xref="GI:12309884"  
/translation="RREYDFAFRLDLCIVYRDGNPY"  
BASE COUNT 18 a 8 c 17 g 23 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,59e-14 Length: 66  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-980-523a-6 (1-22) x AX057042 (1-66)  
QY 1 ArgArgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 1 CGACGTGAGTATATGACTTTCCTTTCCGGGATTATGCATAGTATAGAGATGGGAAT 60  
QY 21 ProTyr 22  
Db 61 CCATAT 66  
RESULT 2  
AF404704 298 bp DNA linear VRL 10-SEP-2001  
LOCUS Human papillomavirus type 16 isolate HPV16E6CC13 E6 protein (E6)  
DEFINITION gene, partial cds.  
ACCESSION AF404704  
VERSION AF404704.1 GI:15529605  
KEYWORDS  
SOURCE Human papillomavirus type 16.  
ORGANISM Human papillomavirus type 16.  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
REFERENCE 1 (bases 1 to 298)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Sequence variation and physical state of human papillomavirus type  
16 cervical cancer isolates from Australia and New Caledonia  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 298)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,  
University of Sydney, Blackburn Building, D06, Off Western Avenue,  
Camperdown, Sydney, New South Wales 2006, Australia  
FEATURES  
source Location/Qualifiers  
1..298  
/organism="Human papillomavirus type 16"  
/isolate="HPV16E6CC13"  
/db\_xref="taxon:10581"  
<1..>298  
/gene="E6"  
CDS  
1..>298  
/gene="E6"  
/codon\_start=2  
/product="E6 protein"  
/protein\_id="AAL01365.1"  
/db\_xref="GI:15529606"  
/translation="DIILECYCKQLLRREVYDFAFRLDLCIVYRDGNPYAVCDKCLK  
FYKISYRHYCYSLYGTTLBQQYKKPLCDLLIRNCQKPLCFEERKORHLDKKQ"  
BASE COUNT 102 a 36 c 64 g 96 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.7e-13 Length: 298  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523a-6 (1-22) x AF404704 (1-298)  
QY 1 ArgArgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 44 CGACGTGAGTATATGACTTTCCTTTCCGGGATTATGCATAGTATAGAGATGGGAAT 103  
QY 21 ProTyr 22  
Db 104 CCATAT 109  
RESULT 3  
S51110 421 bp DNA linear VRL 08-MAY-1993  
LOCUS orf E6 [human papillomavirus HPV, type 16, head and neck tumor,  
Genomic, 421 nt].  
DEFINITION S51110.1 GI:262061  
ACCESSION S51110  
VERSION S51110.1  
KEYWORDS Human papillomavirus head and neck tumor.  
SOURCE Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
ORGANISM Papillomavirus.  
REFERENCE 1 (bases 1 to 421)  
AUTHORS Tyan,Y.S., Liu,S.T., Ong,W.R., Chen,M.L., Shu,C.H. and Chang,Y.S.  
TITLE Detection of Epstein-Barr virus and human papillomavirus in head  
and neck tumors  
JOURNAL J. Clin. Microbiol. 31 (1), 53-56 (1993)  
MEDLINE 93107269  
PUBMED 8380183  
REMARK GenBank staff at the National Library of Medicine created this  
entry [NCBI gibbsq 120826] from the original journal article.  
This sequence comes from Fig. 3B.  
PCR primer sequences: 1-21 and 400-421.  
COMMENT  
FEATURES  
source Location/Qualifiers  
1..421  
/organism="Human papillomavirus"  
/db\_xref="taxon:10566"  
/note="type: 16"  
1..421  
/partial  
gene  
/gene="orf E6"  
BASE COUNT 142 a 63 c 92 g 124 t  
ORIGIN  
Alignment Scores:

Pred. No.: 5.49e-13 Length: 421  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-980-523A-6 (1-22) x S5110 (1-421)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 116 CGACGTGAGGTATATGACCTTTCGCGATTATGATAGTATATAGATGGAAT 175

QY 21 ProTyr 22  
|||||  
Db 176 CCATAT 181

RESULT 4  
AF404696 451 bp DNA linear VRL 10-SEP-2001  
LOCUS Human papillomavirus type 16 isolate HPV16E6CC5 E6 protein (E6) and  
DEFINITION E7 protein (E7) genes, partial cds.  
ACCESSION AF404696  
VERSION AF404696.1 GI:15529581  
KEYWORDS  
SOURCE Human papillomavirus type 16.  
ORGANISM Human papillomavirus type 16  
Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.

REFERENCE 1 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,  
University of Sydney, Blackburn Building, D06, Off Western Avenue,  
Camperdown, Sydney, New South Wales 2006, Australia

FEATURES  
source  
1..451  
/organism="Human papillomavirus type 16"  
/isolate="HPV16E6CC5"  
/db\_xref="taxon:10581"  
gene <1..395  
CDS /gene="E6"  
/gene="E6"  
/codon\_start=3  
/product="E6 protein"  
/protein\_id="AAL01349.1"  
/db\_xref="GI:15529582"  
/translation="TTHNILECVYCKQQLLRREVYDFAFRDLICIVYRDGNPYAVCDK  
CLKFYSKISEYRHVCYSVYGTTLLEQYKNPLCDLLIRINCQKPLCPPEEKQRHLDKKQ  
RHNIRGRWTGRCMSCCRSRTRETQL"  
gene 398..>451  
CDS /gene="E7"  
/gene="E7"  
/codon\_start=1  
/product="E7 protein"  
/protein\_id="AAL01350.1"  
/db\_xref="GI:15529583"  
/translation="MHGDTPTLHEYMLDLQPE"

BASE COUNT 155 a 66 c 96 g 134 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.94e-13 Length: 451  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x AF404696 (1-451)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 54 CGACGTGAGGTATATGACCTTTCGCGATTATGATAGTATATAGATGGAAT 113

QY 21 ProTyr 22  
|||||  
Db 114 CCATAT 119

RESULT 5  
AF404699 451 bp DNA linear VRL 10-SEP-2001  
LOCUS Human papillomavirus type 16 isolate HPV16E6CC8 E6 protein (E6) and  
DEFINITION E7 protein (E7) genes, partial cds.  
ACCESSION AF404699  
VERSION AF404699.1 GI:15529590  
KEYWORDS  
SOURCE Human papillomavirus type 16.  
ORGANISM Human papillomavirus type 16  
Viruses: dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.

REFERENCE 1 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,  
University of Sydney, Blackburn Building, D06, Off Western Avenue,  
Camperdown, Sydney, New South Wales 2006, Australia

FEATURES  
source  
1..451  
/organism="Human papillomavirus type 16"  
/isolate="HPV16E6CC8"  
/db\_xref="taxon:10581"  
gene <1..395  
CDS /gene="E6"  
/gene="E6"  
/codon\_start=3  
/product="E6 protein"  
/protein\_id="AAL01355.1"  
/db\_xref="GI:15529591"  
/translation="TTHNILECVYCKQQLLRREVYDFAFRDLICIVYRDGNPYAVCDK  
CLKFYSKISEYRHVCYSVYGTTLLEQYKNPLCDLLIRINCQKPLCPPEEKQRHLDKKQ  
RHNIRGRWTGRCMSCCRSRTRETQL"  
gene 398..>451  
CDS /gene="E7"  
/gene="E7"  
/codon\_start=1  
/product="E7 protein"  
/protein\_id="AAL01356.1"  
/db\_xref="GI:15529592"  
/translation="MHGDTPTLHEYMLDLHPE"

BASE COUNT 153 a 67 c 97 g 134 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.94e-13 Length: 451  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x AF404699 (1-451)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 54 CGAGCTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATAGATGGGAAT 113

QY 21 ProTyr 22  
|||||  
Db 114 CCATAT 119

RESULT 6  
AF404700 451 bp DNA linear VRL 10-SEP-2001  
LOCUS  
DEFINITION Human papillomavirus type 16 isolate HPV16E6C9 E6 protein (E6) and  
E7 protein (E7) genes, partial cds.  
ACCESSION AF404700  
VERSION AF404700.1 GI:15529593  
KEYWORDS  
SOURCE Human papillomavirus type 16.  
ORGANISM Human papillomavirus type 16  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
REFERENCE 1 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Sequence variation and physical state of human papillomavirus type  
16 cervical cancer isolates from Australia and New Caledonia  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,  
University of Sydney, Blackburn Building, D06, Off Western Avenue,  
Camperdown, Sydney, New South Wales 2006, Australia  
FEATURES  
source  
1..451  
/organism="Human papillomavirus type 16"  
/isolate="HPV16E6C9"  
/db\_xref="taxon:10581"  
<1..395  
/gene="E6"  
<1..395  
/gene="E7"  
/codon\_start=3  
/product="E6 protein"  
/protein\_id="AA01357.1"  
/db\_xref="GI:15529594"  
/translation="TTHDILCEVCYCKQLLRREVYDFAPRDLICIVYRGNPYAVCDK  
CLFYKISIEYRHVCYVGTGLEQQYNKPLCDLLIRINCQKPLCPPEQRHLDDKKQ  
RFNIRGRWTFGRMCCSRSTRRETQL"  
398..>451  
/gene="E7"  
398..>451  
/gene="E7"  
/codon\_start=1  
/product="E7 protein"  
/protein\_id="AA01358.1"  
/db\_xref="GI:15529595"  
/translation="MHGDTPTLHEYMLDLOPE"  
BASE COUNT 155 a 65 c 97 g 134 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.94e-13 Length: 451  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x AF404700 (1-451)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 54 CGAGCTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATAGATGGGAAT 113

QY 21 ProTyr 22  
|||||  
Db 114 CCATAT 119

RESULT 7  
AF404701 451 bp DNA linear VRL 10-SEP-2001  
LOCUS  
DEFINITION Human papillomavirus type 16 isolate HPV16E6C10 E6 protein (E6)  
and E7 protein (E7) genes, partial cds.  
ACCESSION AF404701  
VERSION AF404701.1 GI:15529596  
KEYWORDS  
SOURCE Human papillomavirus type 16.  
ORGANISM Human papillomavirus type 16  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
REFERENCE 1 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Sequence variation and physical state of human papillomavirus type  
16 cervical cancer isolates from Australia and New Caledonia  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 451)  
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,  
University of Sydney, Blackburn Building, D06, Off Western Avenue,  
Camperdown, Sydney, New South Wales 2006, Australia  
FEATURES  
source  
1..451  
/organism="Human papillomavirus type 16"  
/isolate="HPV16E6C10"  
/db\_xref="taxon:10581"  
<1..395  
/gene="E6"  
<1..395  
/gene="E6"  
/codon\_start=3  
/product="E6 protein"  
/protein\_id="AA01359.1"  
/db\_xref="GI:15529597"  
/translation="PIHDIILEVCYCKQLLRREVYDFAPRDLICIVYRGNPYAVCDK  
CLFYKISIEYRHVCYVGTGLEQQYNKPLCDLLIRINCQKPLCPPEQRHLDDKKQ  
RFNIRGRWTFGRMCCSRSTRRETQL"  
398..>451  
/gene="E7"  
398..>451  
/gene="E7"  
/codon\_start=1  
/product="E7 protein"  
/protein\_id="AA01360.1"  
/db\_xref="GI:15529598"  
/translation="MHGDTPTLHEYMLDLOPE"  
BASE COUNT 154 a 66 c 97 g 134 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.94e-13 Length: 451  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x AF404701 (1-451)

QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 54 CGAGCTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATAGATGGGAAT 113

QY 21 ProTyr 22  
|||||  
Db 114 CCATAT 119

```
RESULT 8
AF404703
LOCUS
DEFINITION Human papillomavirus type 16 isolate HPV16E6CC12 E6 protein (E6)
VERSION AF404703
KEYWORDS 451 bp DNA linear VRL 10-SEP-2001
SOURCE Human papillomavirus type 16.
ORGANISM Human papillomavirus type 16.
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts, K.J., Thompson, C.H., Cossart, Y.E. and Rose, B.R.
TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Unpublished
AUTHORS Watts, K.J., Thompson, C.H., Cossart, Y.E. and Rose, B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases, University of Sydney, Blackburn Building, D06, Off Western Avenue, Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source
1. 451
/organism="Human papillomavirus type 16"
/isolate="HPV16E6CC12"
/db_xref="taxon:10581"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AAL01363.1"
/db_xref="GI:15529603"
/translation="TTHNILECYKQQLRRVDFAPRLDCIVYRDGNPYAVCDK
CLFYKISIEYHYCYSLYGTTLQEQYNKPLDLLRCINQKPLCPKQRHLDDKKQ
RPNIRGRWGRGMSCCSSSTRRETQL"
398..>451
/gene="E7"
398..>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AAL01364.1"
/db_xref="GI:15529604"
/translation="MHGDTPTLHEYMLDLQPE"
BASE COUNT 155 a 66 c 95 g 135 t
ORIGIN
Alignment Scores:
Pred. No.: 5.94e-13 Length: 451
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-980-523A-6 (1-22) x AF404703 (1-451)
Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db 54 CGAGCTGAGGTATATGACTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 113
Qy 21 ProTyr 22
Db 114 CCATAT 119
RESULT 9
AR167393
LOCUS
DEFINITION Sequence 29 from patent US 6287569.
VERSION AR167393
KEYWORDS AR167393.1 GI:17903171
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 456)
AUTHORS Kipps, T.J. and Wu, Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 29 11-SEP-2001;
FEATURES
source
1. 456
/organism="unknown"
BASE COUNT 156 a 70 c 101 g 129 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-13 Length: 456
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-980-523A-6 (1-22) x AR167393 (1-456)
Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db 115 CGAGCTGAGGTATATGACTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 174
Qy 21 ProTyr 22
Db 175 CCATAT 180
RESULT 10
AR177943
LOCUS
DEFINITION Sequence 6 from patent US 6313373.
ACCESSION AR177943
VERSION AR177943.1 GI:17920298
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 456)
AUTHORS Eckert, R.L. and Crish, J.F.
TITLE Tissue specific promoters and transgenic mouse for the screening of pharmaceuticals
JOURNAL Patent: US 6313373-A 6 06-NOV-2001;
FEATURES
source
1. 456
/organism="unknown"
BASE COUNT 157 a 72 c 100 g 127 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-13 Length: 456
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-980-523A-6 (1-22) x AR177943 (1-456)
Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db 115 CGAGCTGAGGTATATGACTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 174
Qy 21 ProTyr 22
Db 175 CCATAT 180
```

```
RESULT 11
AF327851
LOCUS
DEFINITION Human papillomavirus type 16 early transforming protein E6 variant
(E6) gene, complete cds.
ACCESSION AF327851
VERSION AF327851.1 GI:12025467
KEYWORDS
SOURCE
ORGANISM Human papillomavirus type 16.
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE
1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, L2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
MEDLINE 96079021
PUBMED 7494284
REFERENCE
2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/isolate="NM 4094"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene
1..456
/gene="E6"
CDS
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AAG45940.1"
/db_xref="GI:12025468"
/translation="MFQDPQPERKPLPCLCTEQTTHDILEVCVCKOQLLRREYD
FAPRDLICIVYRDGNPVAVCDCLKYKSEYRYCYSVYGTLEQYKPLCDLLIR
CINCPPLCEERQRLDKKQRFHNRGWTGRCMCCSSRRRTQL"
BASE COUNT 157 a 71 c 100 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-13 Length: 456
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-980-523A-6 (1-22) x AF327851 (1-456)
QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
|||||
Db 115 CGACGTGAGGTATGACTTTGCTTTTCGGGATTATGCATAGATAGATAGATGGGAAT 174
QY 21 ProTyr 22
|||||
Db 175 CCATAT 180
RESULT 12
HPU34107
LOCUS
DEFINITION Human papillomavirus type 16, isolate NM 4094, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34107
VERSION U34107.1 GI:1098719
KEYWORDS
SOURCE
ORGANISM Human papillomavirus.
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE
1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, L2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
MEDLINE 96079021
PUBMED 7494284
REFERENCE
2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/isolate="NM 4094"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene
1..456
/gene="E6"
CDS
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AAG45940.1"
/db_xref="GI:12025468"
/translation="MFQDPQPERKPLPCLCTEQTTHDILEVCVCKOQLLRREYD
FAPRDLICIVYRDGNPVAVCDCLKYKSEYRYCYSVYGTLEQYKPLCDLLIR
CINCPPLCEERQRLDKKQRFHNRGWTGRCMCCSSRRRTQL"
BASE COUNT 156 a 70 c 101 g 129 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-13 Length: 456
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-980-523A-6 (1-22) x HPU34107 (1-456)
QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
|||||
Db 115 CGACGTGAGGTATGACTTTGCTTTTCGGGATTATGCATAGATAGATAGATGGGAAT 174
QY 21 ProTyr 22
|||||
Db 175 CCATAT 180
RESULT 13
HPU34108
LOCUS
DEFINITION Human papillomavirus type 16, isolate NM 9999, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34108
VERSION U34108.1 GI:1098721
KEYWORDS
SOURCE
ORGANISM Human papillomavirus.
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE
1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, L2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
MEDLINE 96079021
PUBMED 7494284
REFERENCE
2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/isolate="NM 4094"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene
1..456
/gene="E6"
CDS
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AAG45940.1"
/db_xref="GI:12025468"
/translation="MFQDPQPERKPLPCLCTEQTTHDILEVCVCKOQLLRREYD
FAPRDLICIVYRDGNPVAVCDCLKYKSEYRYCYSVYGTLEQYKPLCDLLIR
CINCPPLCEERQRLDKKQRFHNRGWTGRCMCCSSRRRTQL"
BASE COUNT 156 a 70 c 101 g 129 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-13 Length: 456
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-980-523A-6 (1-22) x HPU34107 (1-456)
QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
|||||
Db 115 CGACGTGAGGTATGACTTTGCTTTTCGGGATTATGCATAGATAGATAGATGGGAAT 174
QY 21 ProTyr 22
|||||
Db 175 CCATAT 180
RESULT 14
HPU34109
LOCUS
DEFINITION Human papillomavirus type 16, isolate NM 9999, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34109
VERSION U34109.1 GI:1098722
KEYWORDS
SOURCE
ORGANISM Human papillomavirus.
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
REFERENCE
1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, L2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
MEDLINE 96079021
PUBMED 7494284
REFERENCE
2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/isolate="NM 4094"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene
1..456
/gene="E6"
CDS
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AAG45940.1"
/db_xref="GI:12025468"
/translation="MFQDPQPERKPLPCLCTEQTTHDILEVCVCKOQLLRREYD
FAPRDLICIVYRDGNPVAVCDCLKYKSEYRYCYSVYGTLEQYKPLCDLLIR
CINCPPLCEERQRLDKKQRFHNRGWTGRCMCCSSRRRTQL"
BASE COUNT 156 a 70 c 101 g 129 t
ORIGIN
Alignment Scores:
Pred. No.: 6.01e-13 Length: 456
Score: 124.00 Matches: 22
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-09-980-523A-6 (1-22) x HPU34109 (1-456)
QY 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
|||||
Db 115 CGACGTGAGGTATGACTTTGCTTTTCGGGATTATGCATAGATAGATAGATGGGAAT 174
QY 21 ProTyr 22
|||||
Db 175 CCATAT 180
```



populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments  
J. Virol. 69 (12), 7743-7753 (1995)  
96079021  
MEDLINE  
PUBMED  
7494284  
REFERENCE  
2 (bases 1 to 456)  
Farmer, A.D.  
Direct Submission  
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
Alamos, NM 87501, USA  
Location/Qualifiers  
FEATURES  
source  
1. .456  
/organism="Human papillomavirus"  
/isolate="NM 9999"  
/db\_xref="taxon:10566"  
/note="collected by cervical swab at the University of New Mexico, New Mexico, United States."  
gene  
1. .456  
/gene="E6"  
CDS  
1. .456  
/gene="E6"  
/codon\_start=1  
/product="early transforming protein E6"  
/protein\_id="AAA91655.1"  
/db\_xref="GI:1098722"  
/translation="MFQDPQERPKLPCLCTELQTTIHDIIEVCYCKQQLLRREVVD  
FAPRDLICIVYRDGNPYAVCDKCLKFKYSKISYRYHCYSVYGTTLLEQQYNKPLCDLLIR  
CINCOKPLCPPEEKORHLDDKKORFHNIRGRTGRCMCCRSTRTRETQL"  
BASE COUNT 156 a 70 c 101 g 129 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.01e-13 Length: 456  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x HPU34108 (1-456)  
Qy 1 ArgArgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 115 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGATAGTATATAGAGATGGGAAT 174  
Qy 21 ProTyr 22  
|||||  
Db 175 CCATAT 180  
RESULT 14  
HPU34109  
LOCUS  
DEFINITION Human papillomavirus type 16, isolate NM T197, early transforming protein E6 (E6) gene, complete cds.  
ACCESSION U34109  
VERSION U34109.1 GI:1098723  
KEYWORDS  
SOURCE Human papillomavirus.  
ORGANISM Human papillomavirus  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Yamada, T., Wheeler, C.M., Halpern, A.L., Stewart, A.C., Hildesheim, A. and Jenison, S.A.  
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments  
J. Virol. 69 (12), 7743-7753 (1995)  
96079021  
MEDLINE  
PUBMED  
7494284  
REFERENCE 2 (bases 1 to 456)  
AUTHORS Farmer, A.D.  
TITLE Direct Submission  
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
Alamos, NM 87501, USA  
Location/Qualifiers  
FEATURES  
source  
1. .456  
/organism="Human papillomavirus"  
/isolate="NM T197"  
/db\_xref="taxon:10566"  
/note="collected by cervical swab at the University of New Mexico, New Mexico, United States."  
gene  
1. .456  
/gene="E6"  
CDS  
1. .456  
/gene="E6"  
/codon\_start=1  
/product="early transforming protein E6"  
/protein\_id="AAA91655.1"  
/db\_xref="GI:1098722"  
/translation="MFQDPQERPKLPCLCTELQTTIHDIIEVCYCKQQLLRREVVD  
FAPRDLICIVYRDGNPYAVCDKCLKFKYSKISYRYHCYSVYGTTLLEQQYNKPLCDLLIR  
CINCOKPLCPPEEKORHLDDKKORFHNIRGRTGRCMCCRSTRTRETQL"  
BASE COUNT 156 a 71 c 101 g 128 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.01e-13 Length: 456  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x HPU34109 (1-456)  
Qy 1 ArgArgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 115 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGATAGTATATAGAGATGGGAAT 174  
Qy 21 ProTyr 22  
|||||  
Db 175 CCATAT 180  
RESULT 15  
HPU34110  
LOCUS  
DEFINITION Human papillomavirus type 16, isolate NM T446, early transforming protein E6 (E6) gene, complete cds.  
ACCESSION U34110  
VERSION U34110.1 GI:1098725  
KEYWORDS  
SOURCE Human papillomavirus.  
ORGANISM Human papillomavirus  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Yamada, T., Wheeler, C.M., Halpern, A.L., Stewart, A.C., Hildesheim, A. and Jenison, S.A.  
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments  
J. Virol. 69 (12), 7743-7753 (1995)  
96079021  
MEDLINE  
PUBMED  
7494284  
REFERENCE 2 (bases 1 to 456)  
AUTHORS Farmer, A.D.  
TITLE Direct Submission  
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
Alamos, NM 87501, USA  
Location/Qualifiers  
FEATURES  
source  
1. .456  
/organism="Human papillomavirus"

Direct Submission  
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
Alamos, NM 87501, USA  
Location/Qualifiers  
FEATURES  
source  
1. .456  
/organism="Human papillomavirus"  
/isolate="NM T197"  
/db\_xref="taxon:10566"  
/note="collected by cervical swab at the University of New Mexico, New Mexico, United States."  
gene  
1. .456  
/gene="E6"  
CDS  
1. .456  
/gene="E6"  
/codon\_start=1  
/product="early transforming protein E6"  
/protein\_id="AAA91655.1"  
/db\_xref="GI:1098722"  
/translation="MFQDPQERPKLPCLCTELQTTIHDIIEVCYCKQQLLRREVVD  
FAPRDLICIVYRDGNPYAVCDKCLKFKYSKISYRYHCYSVYGTTLLEQQYNKPLCDLLIR  
CINCOKPLCPPEEKORHLDDKKORFHNIRGRTGRCMCCRSTRTRETQL"  
BASE COUNT 156 a 71 c 101 g 128 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.01e-13 Length: 456  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-980-523A-6 (1-22) x HPU34109 (1-456)  
Qy 1 ArgArgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 115 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGATAGTATATAGAGATGGGAAT 174  
Qy 21 ProTyr 22  
|||||  
Db 175 CCATAT 180  
RESULT 15  
HPU34110  
LOCUS  
DEFINITION Human papillomavirus type 16, isolate NM T446, early transforming protein E6 (E6) gene, complete cds.  
ACCESSION U34110  
VERSION U34110.1 GI:1098725  
KEYWORDS  
SOURCE Human papillomavirus.  
ORGANISM Human papillomavirus  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.  
REFERENCE 1 (bases 1 to 456)  
AUTHORS Yamada, T., Wheeler, C.M., Halpern, A.L., Stewart, A.C., Hildesheim, A. and Jenison, S.A.  
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments  
J. Virol. 69 (12), 7743-7753 (1995)  
96079021  
MEDLINE  
PUBMED  
7494284  
REFERENCE 2 (bases 1 to 456)  
AUTHORS Farmer, A.D.  
TITLE Direct Submission  
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
Alamos, NM 87501, USA  
Location/Qualifiers  
FEATURES  
source  
1. .456  
/organism="Human papillomavirus"

```

/isolate="NM T446"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91657.1"
/db_xref="GI:1098726"
/translation="MFQDPQRPQKLPOLCTELQTHIHIIIECVYCKQQLLRREYVD
FAFRDLCTIVYRDGNPYAVCDKLFYKSIYRHYCYKSVYGTTLQYQYNKPLCDLLIR
CINQKPLCPPEEKQRLDKKQRFHNRGRTGCMSCCRSRTIRRETOL"
BASE COUNT      156 a      71 c      101 g      128 t
ORIGIN

```

Alignment Scores:

Pred. No.:	6.01e-13	Length:	456
Score:	124.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
Db:	14	Gaps:	0

US-09-980-523A-6 (1-22) x HPU34110 (1-456)

```

QY      1  ArgATGGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db      115 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 174

QY      21  ProTyr 22
Db      175 CCATAT 180

```

Search completed: July 6, 2003, 14:07:58  
Job time : 1128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 6, 2003, 12:21:44 ; Search time 39 seconds

(without alignments)  
172.997 Million cell updates/sec

Title: US-09-980-523A-6

Perfect score: 124

Sequence: 1 RREYDFAFRLCIVYRDGNPY 22

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USFTO.spool/US0980523/runat\_03072003\_164406\_4245/app\_query.fasta\_1.199  
-DB=Issued\_Patents\_NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0980523 @CNG 1.1 61 @runat\_03072003\_164406\_4245 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PGTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	456	4	US-09-056-105-29
2	124	100.0	456	4	US-09-430-201-6
3	124	100.0	519	3	US-08-860-165-13
4	124	100.0	519	4	US-09-359-382-13
5	124	100.0	570	5	PCT-US94-05085A-1
6	124	100.0	570	5	PCT-US94-05085-1
7	124	100.0	776	1	US-08-216-233C-1
8	124	100.0	779	4	US-09-210-168-3
9	124	100.0	790	1	US-08-117-083-7
10	124	100.0	801	3	US-08-860-165-9
11	124	100.0	801	4	US-09-359-382-9
12	124	100.0	822	4	US-09-485-885-3

13	124	100.0	879	4	US-09-485-885-9	Sequence 9, Appl
14	124	100.0	1000	1	US-07-965-274-3	Sequence 3, Appl
15	124	100.0	1116	4	US-09-485-885-5	Sequence 5, Appl
16	124	100.0	1173	4	US-09-485-885-13	Sequence 13, Appl
17	124	100.0	7904	1	US-08-316-239B-1	Sequence 1, Appl
18	124	100.0	7904	1	US-08-316-239B-2	Sequence 2, Appl
19	124	100.0	7904	1	US-08-410-005-1	Sequence 1, Appl
20	124	100.0	7904	3	US-08-929-140-1	Sequence 1, Appl
21	124	100.0	7904	4	US-09-560-579A-1	Sequence 1, Appl
22	94	75.8	1095	2	US-08-815-667-12	Sequence 12, Appl
23	94	75.8	1108	2	US-08-815-667-6	Sequence 6, Appl
24	82	66.1	1110	2	US-08-815-667-7	Sequence 7, Appl
25	82	66.1	3283	2	US-08-815-667-14	Sequence 14, Appl
26	79	63.7	7833	1	US-08-074-879-9	Sequence 9, Appl
27	79	63.7	7833	1	US-08-468-057A-9	Sequence 9, Appl
28	78	62.9	1110	2	US-08-815-667-5	Sequence 5, Appl
29	71	57.3	476	2	US-08-486-663A-16	Sequence 16, Appl
30	71	57.3	476	2	US-08-247-904B-9	Sequence 9, Appl
31	71	57.3	476	3	US-08-767-942A-18	Sequence 18, Appl
32	71	57.3	483	5	PCT-US94-05085A-2	Sequence 2, Appl
33	71	57.3	483	5	PCT-US94-05085-2	Sequence 2, Appl
34	71	57.3	817	1	US-08-117-083-11	Sequence 11, Appl
35	71	57.3	837	4	US-09-485-885-20	Sequence 20, Appl
36	71	57.3	1152	4	US-09-485-885-22	Sequence 22, Appl
37	57	46.0	8010	4	US-09-521-526-2	Sequence 2, Appl
38	57	46.0	8010	5	PCT-US95-11859-2	Sequence 2, Appl
39	56	45.2	1107	4	US-09-000-094-19	Sequence 19, Appl
40	56	45.2	1128	4	US-09-000-094-21	Sequence 21, Appl
41	56	45.2	1398	4	US-09-000-094-23	Sequence 23, Appl
42	56	45.2	4770	4	US-09-000-094-45	Sequence 45, Appl
43	55	44.4	1271	1	US-08-474-542A-141	Sequence 141, App
C 44	55	44.4	1271	1	US-08-457-648-141	Sequence 141, App
45	53	42.7	807	3	US-08-718-905-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1

US-09-056-105-29

; Sequence 29 Application US/09056105  
; Patent No. 6287569  
; GENERAL INFORMATION:  
; APPLICANT: KIPPS, THOMAS J.  
; APPLICANT: WU, YUNQI  
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR  
; FILE REFERENCE: 233/221  
; CURRENT APPLICATION NUMBER: US/09/056.105  
; CURRENT FILING DATE: 1998-04-06  
; EARLIER APPLICATION NUMBER: 60/043,467  
; EARLIER FILING DATE: 1997-04-10  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 29  
; LENGTH: 456  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-056-105-29

Alignment Scores:  
Pred. No.: 1.74e-13 Length: 456  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-980-523A-6 (1-22) x US-09-056-105-29 (1-456)

Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
Db 115 CGAGCTGAGGTATATGACTTTGCTTTCGGGATTTATGCTATATATAGATGGGAAT 174

```
QY      21 ProTyr 22
Db      175 CCATAT 180

RESULT 2
US-09-430-201-6
; Sequence 6, Application US/09430201
; Patent No. 6313373
; GENERAL INFORMATION:
; APPLICANT: Eckert, Richard L.
; APPLICANT: Crish, James F.
; TITLE OF INVENTION: tissue Specific Promoters and Transgenic Animals for
; TITLE OF INVENTION: the Screening of Pharmaceuticals
; FILE REFERENCE: CASE-04022
; CURRENT APPLICATION NUMBER: US/09/430,201
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,495
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-430-201-6

Alignment Scores:
Pred. No.:      1-74e-13      Length:      456
Score:          124.00      Matches:      22
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              4      Gaps:          0

US-09-980-523a-6 (1-22) x US-09-430-201-6 (1-456)

QY      1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db      115 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTATGCTAGTATATAGATAGATGGGAAT 174

QY      21 ProTyr 22
Db      175 CCATAT 180

RESULT 3
US-08-860-165-13
; Sequence 13, Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 17227/130
; CURRENT APPLICATION NUMBER: US/08/860,165A
; EARLIER FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(519)
; OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-13

Alignment Scores:
Pred. No.:      2.04e-13      Length:      519
Score:          124.00      Matches:      22
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              4      Gaps:          0

US-09-980-523a-6 (1-22) x US-08-860-165-13 (1-519)

QY      1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db      343 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTATGCTAGTATATAGATAGATGGGAAT 402

QY      21 ProTyr 22
Db      403 CCATAT 408

RESULT 4
US-09-359-382-13
; Sequence 13, Application US/09359382
; Patent No. 6306397
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Stirling John
; APPLICANT: COX, John Cooper
; APPLICANT: WEBB, Elizabeth Ann
; APPLICANT: FRAZER, Ian
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
; FILE REFERENCE: 017227/0148
; CURRENT APPLICATION NUMBER: US/09/359,382
; EARLIER FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 08/860,165
; EARLIER FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: PCT/AU95/00868
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: AU PN0157/94
; EARLIER FILING DATE: 1994-12-20
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Human papillomavirus type 16
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(516)
US-09-359-382-13

Alignment Scores:
Pred. No.:      2.04e-13      Length:      519
Score:          124.00      Matches:      22
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              4      Gaps:          0

US-09-980-523a-6 (1-22) x US-09-359-382-13 (1-519)

QY      1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20
Db      343 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTATGCTAGTATATAGATAGATGGGAAT 402

QY      21 ProTyr 22
Db      403 CCATAT 408

RESULT 5
PCT-US94-05085A-1
; Sequence 1, Application PC/TUS9405085A
; GENERAL INFORMATION:
; APPLICANT: Janice T. Brown
; TITLE OF INVENTION: HUMAN PAPILLOMAVIRUS DETECTION ASSAY
```

NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Baxter Diagnostics Inc.  
STREET: One Baxter Parkway, Building DP-3E  
CITY: Deerfield  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Macintosh System 7.0  
SOFTWARE: Macintosh Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05085A  
FILING DATE: N/A  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/058,920  
FILING DATE: May 6, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark Buonaiuto  
REGISTRATION NUMBER: 31,593  
REFERENCE/DOCKET NUMBER: BA-4448  
TELEPHONE: 708/948-2537  
TELEFAX: 708/948-2642  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
HYPOTHETICAL: no  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Papoviridae, Human papilloma virus  
STRAIN: 16  
FEATURE:  
NAME/KEY: Portion of viral genome coding for E6/E7 polypeptides.  
PUBLICATION INFORMATION:  
AUTHORS: Seedorf, K., Kramer, G., Durst, M.,  
AUTHORS: Suhai, S., and Roweckamp, W.  
TITLE: Human Papillomavirus Type 16 DNA Sequence  
JOURNAL: Virology  
VOLUME: 145  
ISSUE:  
PAGES: 181-185  
DATE: 1985  
PCT-US94-05085A-1

Alignment Scores:  
Pred. No.: 2.29e-13 Length: 570  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
DB: 5 Gaps: 0

US-09-980-523A-6 (1-22) x PCT-US94-05085A-1 (1-570)

Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 137 CGACGTGAGGTATATGACCTTTGCTTTTCGGGATTTATGCATATATATAGATGGGAAT 196  
Qy 21 ProTyr 22  
|||||  
Db 197 CCATAT 202

RESULT 6

PCT-US94-05085-1

; Sequence 1, Application PC/TUS9405085

; GENERAL INFORMATION:

APPLICANT: Janice T. Brown  
TITLE OF INVENTION: HUMAN PAPILLOMAVIRUS DETECTION ASSAY  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Baxter Diagnostics Inc.  
STREET: One Baxter Parkway, Building DP-3E  
CITY: Deerfield  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Macintosh System 7.0  
SOFTWARE: Macintosh Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05085  
FILING DATE: 06-MAY-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/058,920  
FILING DATE: May 6, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark Buonaiuto  
REGISTRATION NUMBER: 31,593  
REFERENCE/DOCKET NUMBER: BA-4448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708/948-2537  
TELEFAX: 708/948-2642  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
HYPOTHETICAL: no  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Papoviridae, Human papilloma virus  
STRAIN: 16  
FEATURE:  
NAME/KEY: Portion of viral genome coding for E6/E7 polypeptides.  
PUBLICATION INFORMATION:  
AUTHORS: Seedorf, K., Kramer, G., Durst, M.,  
AUTHORS: Suhai, S., and Roweckamp, W.  
TITLE: Human Papillomavirus Type 16 DNA Sequence  
JOURNAL: Virology  
VOLUME: 145  
ISSUE:  
PAGES: 181-185  
DATE: 1985  
PCT-US94-05085-1

Alignment Scores:  
Pred. No.: 2.29e-13 Length: 570  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
DB: 5 Gaps: 0

US-09-980-523A-6 (1-22) x PCT-US94-05085-1 (1-570)

Qy 1 ArgArgGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn 20  
|||||  
Db 137 CGACGTGAGGTATATGACCTTTGCTTTTCGGGATTTATGCATATATATAGATGGGAAT 196  
Qy 21 ProTyr 22  
|||||  
Db 197 CCATAT 202

RESULT 7

US-08-216-233C-1

; Sequence 1, Application US/08216233C  
; Patent No. 5506105  
; GENERAL INFORMATION:  
; APPLICANT: HAYDOCK, PAUL V.  
; TITLE OF INVENTION: IN SITU ASSAY OF AMPLIFIED INTRACELLULAR  
; TITLE OF INVENTION: mRNA TARGETS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DADE INTERNATIONAL INC.  
; STREET: 1717 DEERFIELD ROAD  
; CITY: DEERFIELD  
; STATE: ILLINOIS  
; COUNTRY: USA  
; ZIP: 60015  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/216.233C  
; FILING DATE: 22-MAR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 07/808,456  
; FILING DATE: 10-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: TYMESON, CYNTHIA G.  
; REGISTRATION NUMBER: 34,745  
; REFERENCE/DOCKET NUMBER: BA-4203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (305) 222-6423  
; TELEFAX: (305) 222-6686  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 776 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
US-08-216-233C-1  
  
Alignment Scores:  
Pred. No.: 3.34e-13 Length: 776  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0  
  
US-09-980-523A-6 (1-22) x US-08-216-233C-1 (1-776)  
  
QY 1 ArgAtgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
DB 136 CGACGTGAGGTATATGACTTTTCGGGATTTTCATGATATATAGATAGAGATGGGAAT 195  
|||||  
  
QY 21 ProTyr 22  
DB 196 CCATAT 201  
|||||  
  
RESULT 8  
US-09-210-168-3  
; Sequence 3, Application US/09210168  
; Patent No. 6355424  
; GENERAL INFORMATION:  
; APPLICANT: Lorincz, Attila T.  
; TITLE OF INVENTION: ASSESSMENT OF HUMAN PAPILLOMA VIRUS-RELATED DISEASE  
; FILE REFERENCE: 26294005052  
; CURRENT APPLICATION NUMBER: US/09/210,168  
; CURRENT FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: U.S. 60/082,167  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR APPLICATION NUMBER: U.S. 60/070,486

; PRIOR FILING DATE: 1998-01-05  
; PRIOR APPLICATION NUMBER: U.S. 60/069,426  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 779  
; TYPE: DNA  
; ORGANISM: Human papillomavirus  
; FEATURE:  
; OTHER INFORMATION: E6/E7-HPV16  
US-09-210-168-3  
  
Alignment Scores:  
Pred. No.: 3.36e-13 Length: 779  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-980-523A-6 (1-22) x US-09-210-168-3 (1-779)  
  
QY 1 ArgArgGluValTyAspPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
DB 143 CGACGTGAGGTATATGACTTTTCGGGATTTTCATGATATATAGATAGATGGGAAT 202  
|||||  
  
QY 21 ProTyr 22  
DB 203 CCATAT 208  
|||||  
  
RESULT 9  
US-08-117-083-7  
; Sequence 7, Application US/08117083  
; Patent No. 5719054  
; GENERAL INFORMATION:  
; APPLICANT: Boursnell, Michael E.  
; APPLICANT: Inglis, Stephen C.  
; APPLICANT: Munro, Alan J.  
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human  
; TITLE OF INVENTION: Papilloma Virus Proteins  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Walter H. Dreger  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/117,083  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Walter H.  
; REGISTRATION NUMBER: 24,190  
; REFERENCE/DOCKET NUMBER: A-58783  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 790 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

## US-08-117-083-7

## Alignment Scores:

Pred. No.: 3.42e-13 Length: 790  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-980-523A-6 (1-22) x US-08-117-083-7 (1-790)

Qy 1 ArgArgGluValTyArgPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 141 CGACGTGAGGTATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 200

Qy 21 ProTyr 22

|||||

Db 201 CCATAT 206

## RESULT 10

US-08-860-165-9  
; Sequence 9, Application US/08860165A  
; Patent No. 6004557

## GENERAL INFORMATION:

; APPLICANT: EDWARDS, Stirling John  
; APPLICANT: COX, John Cooper  
; APPLICANT: WEBB, Elizabeth Ann  
; APPLICANT: FRAZER, Ian  
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS  
; FILE REFERENCE: 17227/130  
; CURRENT APPLICATION NUMBER: US/08/860.165A  
; EARLIER FILING DATE: 1997-09-22  
; EARLIER APPLICATION NUMBER: PCT/AU95/00868  
; EARLIER FILING DATE: 1995-12-20  
; EARLIER APPLICATION NUMBER: AU PN0157  
; EARLIER FILING DATE: 1994-12-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 801

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(798)

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion

US-08-860-165-9

## Alignment Scores:

Pred. No.: 3.48e-13 Length: 801  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-980-523A-6 (1-22) x US-08-860-165-9 (1-801)

Qy 1 ArgArgGluValTyArgPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 136 CGACGTGAGGTATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 195

Qy 21 ProTyr 22

|||||

Db 196 CCATAT 201

## RESULT 11

US-09-359-382-9  
; Sequence 9, Application US/09359382  
; Patent No. 6306397

## GENERAL INFORMATION:

; APPLICANT: EDWARDS, Stirling John  
; APPLICANT: COX, John Cooper  
; APPLICANT: WEBB, Elizabeth Ann  
; APPLICANT: FRAZER, Ian  
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS  
; FILE REFERENCE: 017227/0148  
; CURRENT APPLICATION NUMBER: US/09/359.382  
; CURRENT FILING DATE: 1999-07-23  
; EARLIER APPLICATION NUMBER: US 08/860.165  
; EARLIER FILING DATE: 1997-09-22  
; EARLIER APPLICATION NUMBER: PCT/AU95/00868  
; EARLIER FILING DATE: 1995-12-20  
; EARLIER APPLICATION NUMBER: AU PN0157/94  
; EARLIER FILING DATE: 1994-12-20  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 801

TYPE: DNA

ORGANISM: Human papillomavirus type 16

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(798)

US-09-359-382-9

## Alignment Scores:

Pred. No.: 3.48e-13 Length: 801  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-980-523A-6 (1-22) x US-09-359-382-9 (1-801)

Qy 1 ArgArgGluValTyArgPheAlaPheArgAspLeuCysIleValTyArgAspGlyAsn 20  
|||||  
Db 136 CGACGTGAGGTATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGAGATGGGAAT 195

Qy 21 ProTyr 22

|||||

Db 196 CCATAT 201

## RESULT 12

US-09-485-885-3  
; Sequence 3, Application US/09485885  
; Patent No. 6342224

## GENERAL INFORMATION:

; APPLICANT: Bruck, Claudine  
; APPLICANT: Cabezon Silva, Teresa  
; APPLICANT: Delisse, Anne-Marie Eva Fernande  
; APPLICANT: Gerard, Catherine Marie Ghislaine  
; APPLICANT: Lombardo-Bencheikh, Angela  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: B45107  
; CURRENT APPLICATION NUMBER: US/09/485.885  
; CURRENT FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/EP98/05285  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: GB 9717953.5  
; PRIOR FILING DATE: 1997-08-22  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3

LENGTH: 822

TYPE: DNA

ORGANISM: Homo sapien

US-09-485-885-3

## Alignment Scores:

Pred. No.: 3.59e-13 Length: 822  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-980-523A-6 (1-22) x US-09-485-885-3 (1-822)

QY 1 ArgArgGluValtyrAspPheAlaPheArgAspLeuCysIleValtyrArgAspGlyAsn 20  
|||||  
Db 454 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGATGGGAAT 513  
QY 21 ProTyr 22  
|||||  
Db 514 CCATAT 519

## RESULT 13

US-09-485-885-9  
; Sequence 9, Application US/09485885

; Patent No. 6342224

; GENERAL INFORMATION:

; APPLICANT: Bruck, Claudine

; APPLICANT: Cabezon Silva, Teresa

; APPLICANT: Delisse, Anne-Marie Eva Fernande

; APPLICANT: Gerard, Catherine Marie Ghislaine

; APPLICANT: Lombardo-Bencheikh, Angela

; TITLE OF INVENTION: Vaccine

; FILE REFERENCE: B45107

; CURRENT APPLICATION NUMBER: US/09/485,885

; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: PCT/EP98/05285

; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: GB 9717953.5

; PRIOR FILING DATE: 1997-08-22

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 9

; LENGTH: 879

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-485-885-9

## Alignment Scores:

Pred. No.: 3,9e-13 Length: 879  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-980-523A-6 (1-22) x US-09-485-885-9 (1-879)

QY 1 ArgArgGluValtyrAspPheAlaPheArgAspLeuCysIleValtyrArgAspGlyAsn 20  
|||||  
Db 511 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGATGGGAAT 570  
QY 21 ProTyr 22  
|||||  
Db 571 CCATAT 576

## RESULT 14

US-07-965-274-3

; Sequence 3, Application US/07965274

; Patent No. 5501947

; GENERAL INFORMATION:

; APPLICANT: EMERY, VINCENT C.

; APPLICANT: BAVIN, PEGGY J.

; APPLICANT: WALKER, PATRICK

; TITLE OF INVENTION: METHOD FOR THE DIAGNOSIS OF HUMAN

; TITLE OF INVENTION: PAPILLOMA VIRUS TYPE 16 USING THE POLYMERASE CHAIN

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: REED & ROBINS

; STREET: 635 BRYANT STREET

; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/965,274  
; FILING DATE: 12-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB91/01212  
; FILING DATE: 19-JUL-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9015845.2  
; FILING DATE: 19-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5150-0028  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 617-8999  
; TELEFAX: (415) 327-3231  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1000 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-07-965-274-3

## Alignment Scores:

Pred. No.: 4.57e-13 Length: 1000  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-980-523A-6 (1-22) x US-07-965-274-3 (1-1000)

QY 1 ArgArgGluValtyrAspPheAlaPheArgAspLeuCysIleValtyrArgAspGlyAsn 20  
|||||  
Db 218 CGACGTGAGGTATATGACTTTGCTTTTCGGGATTTATGCATAGTATATAGATGGGAAT 277  
QY 21 ProTyr 22  
|||||  
Db 278 CCATAT 283

## RESULT 15

US-09-485-885-5

; Sequence 5, Application US/09485885

; Patent No. 6342224

; GENERAL INFORMATION:

; APPLICANT: Bruck, Claudine

; APPLICANT: Cabezon Silva, Teresa

; APPLICANT: Delisse, Anne-Marie Eva Fernande

; APPLICANT: Gerard, Catherine Marie Ghislaine

; APPLICANT: Lombardo-Bencheikh, Angela

; TITLE OF INVENTION: Vaccine

; FILE REFERENCE: B45107

; CURRENT APPLICATION NUMBER: US/09/485,885

; CURRENT FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: PCT/EP98/05285

; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: GB 9717953.5

; PRIOR FILING DATE: 1997-08-22

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: FastSeq for Windows Version 3.0



; SEQ ID NO 5  
; LENGTH: 1116  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-485-885-5

Alignment Scores:  
Pred. No.: 5.23e-13 Length: 1116  
Score: 124.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-980-523A-6 (1-22) x US-09-485-885-5 (1-1116)

QY	1	ATGATGluValTyrAspPheAlaPheArgAspLeuCysIleValTyrArgAspGlyAsn	20
DB	454	CGACGTGAGGTATATGACITTCGCTTTTCGGGATTATGCATATATATAGATGGGAAT	513
QY	21	ProTyr 22	
DB	514	CCATAT 519	

Search completed: July 6, 2003, 13:08:12  
Job time : 41 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 16:39:56 ; Search time 40 Seconds  
(without alignments)  
52.874 Million cell updates/sec

Title: US-09-980-523a-6

Perfect score: 124

Sequence: 1 RREYDFAFRDLCIVYRDGNPY 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	158	1 W6WL35	protein E6 - human
2	93	75.0	149	1 W6WL35	E6 protein - human
3	93	75.0	149	1 W6WL33	E6 protein - human
4	92	74.2	149	1 W6WL58	E6 protein - human
5	90	72.6	148	2 S36515	E6 protein - human
6	88	71.0	148	2 A61237	E6 protein - human
7	88	71.0	148	2 S36573	E6 protein - human
8	83	66.9	158	2 S36561	E6 protein - human
9	79	63.7	154	2 S36527	E6 protein - human
10	79	63.7	158	1 W6WL39	E6 protein - human
11	78	62.9	158	1 W6WLPR	E6 protein - human
12	73	58.9	151	1 W6WL51	E6 protein - human
13	72	58.1	149	1 W6WL31	E6 protein - human
14	71	57.3	158	1 W6WL18	E6 protein - human
15	70	56.5	155	1 W6WL43	E6 protein - human
16	68	54.8	153	2 S36503	E6 protein - human
17	68	54.8	191	1 W6WLR1	E6 protein - human
18	66	53.2	150	2 S36544	E6 protein - human
19	65	52.4	155	1 W6WL56	E6 protein - human
20	58	46.8	150	1 W6WL44	E6 protein - human
21	56	45.2	150	1 W6WL6	E6 protein - human
22	56	45.2	154	2 S36555	E6 protein - human
23	54	43.5	150	1 W6WL11	E6 protein - human
24	53	42.7	142	2 S36509	E6 protein - human
25	52	41.9	154	2 S36584	E6 protein - human
26	51	41.1	152	2 S36550	E6 protein - human
27	51	41.1	517	2 T37225	hypothetical prote
28	50	40.3	322	2 B69690	ribose ABC transpo
29	49	39.5	148	2 S36532	E6 protein - human

#### ALIGNMENTS

##### RESULT 1

W6WL35

protein E6 - human papillomavirus type 16

C:Species: human papillomavirus type 16

C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 20-Aug-1999

C:Accession: A03682; T10427

R:Seedorf, K.; Kramer, G.; Durst, M.; Suhai, S.; Rowekamp, W.G.

Virology 145:160-165, 1985

A:Title: Human papillomavirus type 16 DNA sequence.

A:Reference number: A22355; MUID:85246220; PMID:2990099

A:Accession: A03682

A:Molecule type: DNA

A:Residues: 1-158 <SEE>

A:Cross-references: GB:K02718; NID:g333031; PIDN:AAA46939.1; PID:g333032

R:Kennedy, I.M.; Haddow, J.K.; Clements, J.B.

J. Virol. 65, 2093-2097, 1991

A:Title: A negative element in the human poapillomavirus type 16 genome acts at the 1

A:Reference number: Z17014; MUID:91162763; PMID:1848319

A:Accession: T10427

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-158 <REN>

A:Cross-references: EMBL:K02718; NID:g333031; PIDN:AAA46939.1; PID:g333032

A:Gene: E6

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; zinc finger

F;37-73/Region: zinc finger CCCC motif

F;110-146/Region: zinc finger CCCC motif

Query Match 100.0%; Score 124; DB 1; Length 158;

Best Local Similarity 100.0%; Pred. No. 8.3e-12; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RREYDFAFRDLCIVYRDGNPY 22

Db 46 RREYDFAFRDLCIVYRDGNPY 67

##### RESULT 2

W6WL35

E6 protein - human papillomavirus type 35

C:Species: human papillomavirus type 35

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Mar-2001

C:Accession: E40824; S36521

R:Marich, J.E.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.

Virology 186, 770-776, 1992

A:Title: The phylogenetic relationship and complete nucleotide sequence of human pap

A:Reference number: A40824; MUID:92124753; PMID:1310198

A:Accession: E40824

A>Status: translation not shown

```

F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif

Query Match          74.2%; Score 92; DB 1; Length 149;
Best Local Similarity 77.3%; Pred. No. 6.8e-07;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps

QY 1 RREYDYFAFRDLCIVYRDGNPY 22
    | | | | | | | | | | | | | |
Db 39 RSEYDYFVFDLRIYVRDGNPF 60

RESULT 5
S36515
E6 protein - human papillomavirus type 34
C:Species: human papillomavirus type 34
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36515
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36515
A:Molecule type: DNA
A:Residues: 1-148 <DEL>
A:Cross-references: EMBL:X74476; NID:g396989; PIDN:CAA52555.1; PID:g996990
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match          72.6%; Score 90; DB 2; Length 148;
Best Local Similarity 76.2%; Pred. No. 1.4e-06;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps

QY 1 RREYDYFAFRDLCIVYRDGNP 21
    | | | | | | | | | | | | | |
Db 40 RCEYDYFIFRDLVYVRGKRP 60

RESULT 6
A61237
E6 protein - human papillomavirus type 52
C:Species: human papillomavirus type 52
C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 03-May-1996
C:Accession: A61237
R:Takami, Y.; Kondoh, G.; Saito, J.; Noda, K.; Sudiro, T.M.; Sjahurachman,
Int. J. Cancer 48, 516-522, 1991
A:Title: Cloning and characterization of human papillomavirus type 52 from a
A:Reference number: A61237; MUID:91258022; PMID:1646174
A:Accession: A61237
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <TAK>
C:Superfamily: papillomavirus E6 protein

Query Match          71.0%; Score 88; DB 2; Length 148;
Best Local Similarity 77.3%; Pred. No. 2.8e-06;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps

QY 1 RREYDYFAFRDLCIVYRDGNPY 22
    | | | | | | | | | | | | | |
Db 39 RREYKFLFDLRIYVRDNNPY 60

RESULT 7
S36573
E6 protein - human papillomavirus type 52
C:Species: human papillomavirus type 52
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36573
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469

```

A:Accession: S36573  
A:Molecule type: DNA  
A:Residues: 1-148 <DEL>  
A:Cross-references: EMBL:X74481; NID:g397038; PIDN:CAA52585.1; PID:g397039  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 71.0%; Score 88; DB 2; Length 148;  
Best Local Similarity 77.3%; Pred. No. 2.8e-06;  
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCTIVRDGNPY 22  
||||| | | | | | | | | |  
Db 39 RREYKFLFTDLRIVYRDNNPY 60

## RESULT 8

S36561 E6 protein - human papillomavirus type 45

C:Species: human papillomavirus type 45

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36561

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36561

A:Molecule type: DNA

A:Residues: 1-158 <DEL>

A:Cross-references: EMBL:X74479; NID:g397022; PIDN:CAA52573.1; PID:g397023

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 66.9%; Score 83; DB 2; Length 158;  
Best Local Similarity 83.3%; Pred. No. 1.8e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCTIVYRD 18  
| | | | | | | | | | | | | |  
Db 41 RTEYQFAFRDLCTIVYRD 58

## RESULT 9

S36527 E6 protein - human papillomavirus type 53

C:Species: human papillomavirus type 53

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36527

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36527

A:Molecule type: DNA

A:Residues: 1-154 <DEL>

A:Cross-references: EMBL:X74482; NID:g397046; PIDN:CAA52591.1; PID:g397047

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 63.7%; Score 79; DB 2; Length 154;  
Best Local Similarity 70.0%; Pred. No. 7.2e-05;  
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EVYDFAFRDLCTIVYRDGNY 22  
| | | | | | | | | | | | | |  
Db 45 EVYNFAVYTLRVYRDGYPY 64

## RESULT 10

W6WL39

E6 protein - human papillomavirus type 39

C:Species: human papillomavirus type 39

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
C:Accession: A38502  
R:Volpers, C.; Strecek, R.E.  
Virology 181, 419-423, 1991  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 39.  
A:Reference number: A38502; MUID:91135017; PMID:1847266  
A:Accession: A38502  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-158 <VOL>  
A:Cross-references: GB:M62849; EMBL:M38185; NID:g333245; PIDN:AAA47050.1; PID:g463186  
C:Superfamily: papillomavirus E6 protein  
C:Keywords: DNA binding; early protein; transforming protein; zinc finger  
F:32-68/Region: zinc finger CCCC motif  
F:105-141/Region: zinc finger CCCC motif

Query Match 63.7%; Score 79; DB 1; Length 158;  
Best Local Similarity 73.7%; Pred. No. 7.4e-05;  
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EVYDFAFRDLCTIVYRDGNP 21  
| | | | | | | | | | | | | |  
Db 43 EVTEFAFSDLYVYVRDGE 61

## RESULT 11

W6WLP9

E6 protein - human papillomavirus type ME180 (provirus)

C:Species: human papillomavirus type ME180

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Feb-1997

C:Accession: C40509

R:Reuter, S.; Delius, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwarz, E.

J. Virol. 65, 5564-5568, 1991

A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma

A:Reference number: A40509; MUID:91374616; PMID:1716694

A:Accession: C40509

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-158 <REU>

A:Cross-references: GB:M73258

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; zinc finger

F:32-68/Region: zinc finger CCCC motif

F:105-141/Region: zinc finger CCCC motif

Query Match 62.9%; Score 78; DB 1; Length 158;  
Best Local Similarity 71.4%; Pred. No. 0.00011;  
Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RREYDFAFRDLCTIVYRDGNP 21  
| | | | | | | | | | | | | |  
Db 41 RTEYIEFAFGDLNVYVRDGV 61

## RESULT 12

W6WL51

E6 protein - human papillomavirus type 51

C:Species: human papillomavirus type 51

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997

C:Accession: E40415

R:Lungu, O.; Crum, C.P.; Silverstein, S.J.

J. Virol. 65, 4216-4225, 1991

A:Title: Biologic properties and nucleotide sequence analysis of human papillomavirus

A:Reference number: A40415; MUID:91303675; PMID:1649326

A:Accession: E40415

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-151 <LUN>

A:Cross-references: GB:M62877

C:Superfamily: papillomavirus E6 protein

C:Keywords: DNA binding; early protein; zinc finger

Search completed: July 3, 2003, 16:42:02  
Job time : 41 secs